RAND

Analyses for the Initial Implementation of the Inpatient Rehabilitation Facility Prospective Payment System

Grace M. Carter, Melinda Beeuwkes Buntin, Orla Hayden, Jennifer Kawata, Susan M. Paddock, Daniel A. Relles, Gregory K. Ridgeway, Mark E. Totten, Barbara O. Wynn

RAND Health

20020805 175

RAND

Analyses for the Initial Implementation of the Inpatient Rehabilitation Facility Prospective Payment System

Grace M. Carter, Melinda Beeuwkes Buntin, Orla Hayden, Jennifer Kawata, Susan M. Paddock, Daniel A. Relles, Gregory K. Ridgeway, Mark E. Totten, Barbara O. Wynn

RAND Health

MR-1500-CMS

A profile of RAND Health, abstracts of its publications, and ordering information can be found on the RAND Health home page at www.rand.org/health.

RAND is a nonprofit institution that helps improve policy and decisionmaking through research and analysis. RAND® is a registered trademark. RAND's publications do not necessarily reflect the opinions or policies of its research sponsors.

PREFACE

This report describes the research that RAND performed to support the efforts of the Health Care Financing Administration (HCFA) to design, develop, and implement a Prospective Payment System (PPS) for inpatient rehabilitation. It presents recommendations concerning the payment system and discusses our plans for further research on the monitoring and refinement of the PPS.

In the Balanced Budget Act of 1997, Congress mandated that HCFA implement a PPS for inpatient rehabilitation. The Centers for Medicare and Medicaid Services (CMS, the successor agency to HCFA) issued the final rule governing such a PPS on August 7, 2001. The PPS started on January 1, 2002. This rule reflects policy analyses and decisions that took into account other contextual analyses and the experience of other PPSs, as well as the research presented here. This report is a research document, not a policy document, and should be viewed as a contribution to understanding inpatient rehabilitation.

The final report of our project consists of this volume and four others:

Carter, G. M., M. Beeuwkes-Buntin, O. Hayden, S. M. Paddock, J. Kawata, D. A. Relles, G. K. Ridgeway, M. E. Totten, and B. O. Wynn (2002). *Executive Summary of Analyses for the Initial Implementation of the Inpatient Rehabilitation Facility Prospective Payment System.* Santa Monica, CA: RAND, MR-1500/1-CMS.

Buchanan, J. L., P. Andres, S. Haley, S. M. Paddock, D. C. Young, and A. Zaslavsky (forthcoming). *Final Report on Assessment Instruments for PPS.* Santa Monica, CA: RAND, MR-1501-CMS.

Relles, D. A., and G. M. Carter (forthcoming). *Linking Medicare and Rehabilitation Hospital Records to Support Development of a Rehabilitation Hospital Prospective Payment System.* Santa Monica, CA: RAND, MR-1502-CMS.

Carter, G. M., D. A. Relles, B. O. Wynn, J. Kawata, S. M. Paddock, N. Sood, and M. E. Totten (2000). *Interim Report on an Inpatient Rehabilitation Facility Prospective Payment System.* Santa Monica, CA: RAND, DRU-2309 (forthcoming as MR-1503-CMS).

The first volume listed above summarizes the major findings from all these reports. The second volume evaluates two alternative data instruments that had been proposed as a means of gathering case mix data from hospitals. Relles and Carter (2002) provides additional information on our database, especially about the methodology used to link multiple sources of data describing the same discharge. The last volume is a preliminary report that HCFA used to help prepare the Notice of Proposed Rule Making during the summer and fall of 2000.

Together, these documents constitute the final report for phase I of our study of the design, development, monitoring, and refining of an Inpatient Rehabilitation Facility Prospective Payment System.

CONTENTS

FIGURES

TABLES

ACKNOWLEDGMENTS

Members of Technical Expert Panel on Inpatient Rehabilitation
Facility Prospective Payment System

Ken Aitchison
Kessler Institute for Rehabilitation

Joan August
Cedars-Sinai Medical Center

James Ball
Catholic Health Services

Jean Davis
HealthSouth Corp.

Susan Dean-Baar
University of Wisconsin

Frank Gainer
Greater Southeast Community Hospital

Norbert Goldfield
3MHIS

Stuart Guterman
Congressional Budget Office

Kurt Hoppe
Iowa Health System

Brad Hutchins
Oral Health America

Alan Jette
Boston University

Robert Kane
University of Minnesota

Sally Kaplan
MedPAC

Richard Linn
State University of New York

John Melvin
Moss Rehabilitation Group

John Morris
Hebrew Rehab. Center for Aged
Research & Training Institute

Paul Rao
National Rehabilitation Hospital

Pam Roberts
Cedars-Sinai Medical Center

Elliot Roth
Northwestern University Medical
School

Barry Smith
Baylor Health Systems

Margaret Stineman
University of Pennsylvania

Carolyn Zollar
AMRPA

ACRONYMS

ADC	Average daily census
BBA	Balanced Budget Act of 1997
BBRA	Balanced Budget Refinement Act of 1999
BI	Brain injury
CART	Classification and regression trees
CC	Complication or comorbidity
CCR	Cost-to-charge ratio
CF	Conversion factor
CMG	Case mix group
CMI	Case mix index
CMS	Centers for Medicare and Medicaid Services (successor agency to HCFA)
COS	Clinical Outcomes Systems
CY	Calendar year
DRG	Diagnosis-Related Group
ESRD	End-stage renal disease
FIM	Functional Independence Measure
FRG	Function Related Group
FRGC	Function Related Group with Comorbidities
FTE	Full-time equivalent
FY	Fiscal year
GAM	Generalized additive models
HCFA	Health Care Financing Administration
HCRIS	Hospital Cost Report Information System
HHA	Home health agency
HMO	Health Maintenance Organization
HSRV	Hospital specific relative value
IRF	Inpatient Rehabilitation Facility
LE	Lower extremity
LOS	Length of stay
LTC	Long-term care
MART	Multiple adaptive regression trees
MDC	Major Diagnostic Category

MDS-PAC	Minimum Data Set-Post Acute Care
MEDPAR	Medicare Provider Analysis and Review
MMT	Major multiple trauma
MSA	Metropolitan Statistical Area
NH	Nursing home
NOS	Not otherwise specified
NPRM	Notice of Proposed Rule Making
NTSCI	Nontraumatic spinal cord injury
OACT	CMS Office of the Actuary
OLS	Ordinary least squares
PAC	Post-acute care
PAI	Patient assessment instrument
PPS	Prospective Payment System
PTC	Payment-to-cost [ratio]
RIC	Rehabilitation Impairment Category
RMSE	Root mean square error
RRC	Residency Review Committee
SCI	Spinal cord injury
SEE	Standard error of estimate
SNF	Skilled nursing facility
SSI	Supplemental Security Income
TBI	Traumatic brain injury
TEFRA	Tax Equity and Fiscal Responsibility Act
TEP	Technical Expert Panel
TSCI	Traumatic spinal cord injury
UDSmr	Uniform Data System for medical rehabilitation
UTI	Urinary tract infection
WI	Wage index

1. INTRODUCTION

BACKGROUND

In the Balanced Budget Act of 1997, Congress mandated that the Health Care Financing Administration (HCFA) implement a Prospective Payment System (PPS) for inpatient rehabilitation under Medicare. This new PPS was implemented beginning on January 1, 2002. The Centers for Medicare and Medicaid Services (CMS, the successor agency to HCFA) issued the final rule governing the PPS on August 7, 2001. This report describes the research that RAND performed to support HCFA's efforts to design, develop, and implement this PPS. It presents recommendations concerning the payment system and discusses our plans for further research on the monitoring and refinement of the PPS.

The initial design of the system was first presented in a Notice of Proposed Rule Making (NPRM) (HCFA, 2000). Our interim report, Carter et al. (2000), presented analyses that HCFA used to help make its decisions in the NPRM. In this report, we update these analyses using later data. We also improve the analysis and our recommendations to HCFA by taking into account comments made by our Technical Expert Panel in its review of our interim report. This is a report of research. The final decisions made by CMS and the rationale for those decisions may be found in the rule governing the IRF PPS (CMS, 2001).

The new PPS applies to rehabilitation hospitals and to distinct rehabilitation units of acute care hospitals, which are excluded from the acute care PPS. To qualify for such exclusion, rehabilitation facilities must meet two conditions. First, Medicare patients must receive intensive therapy (generally at least three hours per day). Second, 75 percent of each facility's patients must have one of 10 specified problems related to neurological or musculoskeletal disorders or burns. We call this PPS the Inpatient Rehabilitation Facility PPS, or IRF PPS.

Need to Improve Payment

Payment for inpatient care of Medicare beneficiaries in a rehabilitation facility was--and for many facilities still is, in whole

or in part--made under the Tax Equity and Fiscal Responsibility Act
(TEFRA) of 1982. The payment amount depends on a per-case target amount
that is calculated from historical costs at the facility trended forward
and on the hospital's actual cost per case. Under TEFRA, there is no
adjustment for changes in a hospital's case mix, and new hospitals were
able to obtain larger payments than existing hospitals by spending more
on care during their base years. The Balanced Budget Act of 1997 (BBA)
introduced interim changes to the payment system designed to reduce
HCFA's costs and to mitigate the advantage that new hospitals had under
the TEFRA payment system. In particular, limits were set on the payment
rate for new hospitals, separate maximum payment limits for all
hospitals were created, and future update rates were greater for
hospitals whose costs exceeded payments. In addition, hospitals that
were receiving Medicare payments prior to FY 1990 were allowed to
request rebasing of their target amounts.

Technological changes in the process of care, greater availability
of post-acute care, and financial incentives for acute care hospitals to
release patients quickly combined to cause rapid growth in Medicare
payments for all forms of post-acute care, including rehabilitation. The
number of Medicare beneficiaries served by skilled nursing facilities
(SNFs) grew by 94 percent from 1990 to 1995, the number served by home
health agencies (HHAs) grew by 78 percent, and the number of Medicare
discharges from rehabilitation facilities grew by 67 percent (MedPAC,
1998, Charts 4-3, 4-8, and 4-17). Acute care hospitals, paid under the
acute PPS, found it advantageous to transfer patients to a different
setting as soon as possible. Probably affected by both PPS and TEFRA
incentives, the number of rehabilitation hospitals and units increased
4.1 percent annually from 1990 to 1997 (MedPAC, 1998). By FY 1997, 25.3
percent of acute care PPS discharges used post-acute care within one day
of discharge and 2.9 percent went to a rehabilitation facility (MedPAC,
2001).

Although rehabilitation facility payments from Medicare were
substantially less than costs in the early 1990s, the ratio of aggregate
Medicare payments to cost increased rapidly during the decade. By 1995,
payments exceeded costs by 7 percent in freestanding rehabilitation
hospitals and by 4 percent in rehabilitation units (MedPAC, 1998, Chart

4.17). This improved position was driven, at least in part, by reduced costs associated with a decline in length of stay (LOS) for rehabilitation patients.

In addition to TEFRA's inability to control Medicare expenditures, it also may hinder access to care. The lack of a case mix adjustment in TEFRA creates incentives for providers to specialize in relatively less-expensive cases, which could conceivably limit beneficiary access. Further, TEFRA lacks outlier payments, which help to mitigate the acute PPS's incentives to underserve the most expensive patients and that provide substantial protection to providers against financial risk (Keeler, Carter, and Trude, 1988). Additional distortion of case-level payments occurs when TEFRA counts discharges that do not include a full course of rehabilitation (e.g., short stays for evaluation, transfer cases) as full cases. These distortions may have both quality and cost control implications.

TEFRA was widely perceived to be unfair to older hospitals. Until the Balanced Budget Act of 1997, newer hospitals were not subject to the same incentives for efficiency, and indeed were rewarded for incurring higher costs in their base year(s).

Research Enabling an IRF PPS

One of the reasons for the initial exclusion of rehabilitation hospitals from the PPS was that Diagnosis-Related Groups (DRGs) could not predict resource use at these facilities very well. Functional status, measured by activities of daily living and mobility, is more correlated with patient charges than are diagnoses (Hosek et al., 1986). Because restoring functional status is the goal of rehabilitation, functional status at admission is one of the primary determinants of resource use.

In the early 1990s, Margaret Stineman and colleagues developed Function Related Groups (FRGs) based on the Functional Independence Measure (FIM) and on a clinically derived set of rehabilitation impairment categories (RICs). The FIM is an 18-item measure covering six domains: self-care, sphincter control, mobility, locomotion, communication, and social cognition (Stineman, Hamilton, et al., 1994).

The response to each item on the FIM ranges from 1 (least independent) to 7 (most independent).

Carter, Relles, and Buchanan (1997) evaluated the FIM-FRGs and found that they use the correct organizing concepts for a rehabilitation patient classification system: impairment groups subdivided by functional status and age. The study found further that FIM-FRGs are good predictors of resource use. The analysis suggested that the FIM-FRGs could be a suitable basis for a rehabilitation PPS, but that certain modifications would produce even better groups for payment purposes. In particular, the authors advised using a multiplicative factor to account for the extra costs associated with patients who have at least one of a selected set of comorbidities. They also expanded the FRG set to 82 FRGs. In expanding the number of FRGs, they changed the algorithm to reduce the number of categories that it would produce. Carter, Buchanan, et al. (1997) described the construction of a model of a rehabilitation PPS based on these expanded FRGs and comorbidity weights. They examined the major elements of such a system: case weights, payment arrangements for unusual cases such as transfers and outliers, hospital-level payment adjustments, and a monitoring system. They examined alternative forms of several of these payment elements in payment simulations. They concluded that a PPS based on the FIM-FRGs is feasible and could achieve several goals. They judged that it would

- provide hospitals with incentives for efficiency because they can keep payments in excess of costs;
- promote access for all Medicare beneficiaries to high-quality and appropriate care because the system pays appropriately for clinical or demographic sub-groups;
- be fair to hospitals because it
  - distributes Medicare payments according to patient characteristics modified by input prices and
  - covers costs at all groups of hospitals except those that probably had high costs because their payments were especially high under the payment system in use at that time.

ASSESSMENT INSTRUMENT FOR THE IRF PPS

The assessment instrument used to determine case classification for payment under the IRF PPS is called the Inpatient Rehabilitation

Facility Patient Assessment Instrument (IRF PAI). It includes the 18 functional items that are part of the FIM instrument. Each item is an assessment of how independently the patient can accomplish 18 simple activities: six self-care items (eating, grooming, bathing, dressing upper body, dressing lower body, and toileting), two sphincter control items (bowel and bladder), five items on transfer and locomotion (transfers to and from bed/chair/wheelchair, toilet, tub/shower, locomotion by walking or wheelchair, and stairs), and five items on communication and cognition (comprehension, expression, social interaction, problem solving, and memory). The IRF PAI also asks for the primary reason for admission to the rehabilitation program.

In the NPRM, HCFA had proposed using a new instrument called the Minimum Data Set--Post-Acute Care, or MDS-PAC. This instrument was developed with the intent that it be used for all post-acute settings, although it differed substantially from any instrument actually in use in other settings. The decision to create a new assessment instrument that would include the FIM was made after a study by Buchanan et al. (forthcoming) showed that the MDS-PAC could not be used to reliably assign the case mix groups (CMGs) defined in the NPRM. The study also showed that the administrative burden of the MDS-PAC was significantly greater than that for the FIM. It took an average of 147 minutes for the hospital staff to fill out the MDS-PAC compared to only 25 minutes for the FIM.[1]

The Buchanan et al. study also showed that many individual FIM items were less reliable than one would want in a payment instrument. The FIM itself is more complex than the 18 items appear to be from the survey instrument. For example, to answer the bladder control item, one must first answer two related questions; the minimum of the two responses is then used. In order to improve the reliability of the FIM items, the new IRF PAI asks hospitals to record each of the FIM sub-items. Another way that the IRF PAI questions differ from the FIM is that they ascertain whether or not the activity being rated was actually

---

[1] Since all hospital teams routinely use the FIM, some of the discrepancy in time is likely due to start-up effects. Nevertheless, it is almost certain that the MDS-PAC would take much longer than the FIM to fill out even after familiarity has been achieved.

observed. The FIM instructions are to record "least independent" when an item's activity is not observed. The new format should improve the comparability of information across hospitals.

PAYMENT UNDER THE IRF PPS

The unit of payment in the IRF PPS is a Medicare-covered hospital stay, beginning with admission to the rehabilitation hospital or unit and ending with discharge from that facility. Inpatient rehabilitation is inherently episodic: Episodes typically begin with a clinical event leading to acute care, and the majority end with a return to independent living in the community. Indeed, return to the community is the stated goal of the inpatient rehabilitation process. The Balanced Budget Refinement Act of 1999 (BBRA) mandated that discharges be the unit of payment.

The IRF PPS is used for payment of discharges that occur after the start of the facility's cost reporting period that begins on or after January 1, 2002. The hospital's payment rate is a blend of two-thirds of the national IRF PPS payment and one-third of its TEFRA payment, although each facility can opt to be paid at 100 percent of the PPS rate instead of the blend. A fully national prospective payment will be used for all cost report periods beginning in FY 2003. The following formula describes the calculation of the IRF PPS payment for each case.

Each case is classified into a case mix group. Almost all CMGs are based on impairment, functional status as measured by items from the FIM, and comorbidities. Additional groups were constructed for deaths and atypical short-stay cases whose resource use is not well described by these characteristics. The CMGs are assigned based on information in the IRF PAI.

The IRF PPS payment for a discharge in hospital i in CMG k is given by

$$F = R * A_i * W_k ,$$

where R is the national conversion factor, A_i is the facility payment adjustment, and W_k is the CMG relative weight. In FY 2002, R was chosen to meet the statutory budget neutrality constraint that payment under the new PPS equal what payment would have been under TEFRA, as estimated by CMS's Office of the Actuary (OACT).

This payment may be increased by outlier payments. Also, short-stay transfer cases receive a payment for each day in the hospital plus a case-level payment equal to one-half of one day's payment.

OVERVIEW OF THE REPORT

Approach

We examined a variety of options for the elements of the IRF PPS, such as the classification system and the facility payment adjustment, and we analyzed the distribution of funds under each payment option. The criteria for the design and development of the IRF PPS are similar to those used in Carter, Buchanan, et al. (1997). To insure access to quality care for all Medicare patients, the system must identify groups of patients who need different levels of resources and then pay for each group in proportion to cost. The system should be fair to hospitals by paying for costs that are outside the control of hospital administrators, such as area wage levels or a population that is disproportionately poor. The payment system must also allow CMS to control its budget for post-acute care. It must provide incentives for hospitals to provide quality care and limit incentives for "gaming the system."

In order to meet these criteria, all major system parameters are based on data on rehabilitation facilities' case mix and cost. We began our work using only 1996 and 1997 data and made recommendations to HCFA for its use in developing the NPRM. These analyses and recommendations are recorded in our interim report (Carter et al., 2000). That report was reviewed by a Technical Expert Panel (TEP), which made several very helpful suggestions for improving the research. In response to suggestions from CMS based in part on public comment on the NPRM and from our TEP, we extended our research and updated it to include data from CY 1998 and CY 1999.

The research leading to the implementation of the system in FY 2002 is the first phase of our work. CMS is committed to developing a system to monitor the IRF PPS, update parameters, and refine the IRF PPS to better meet its goals. The second phase of our work will be to help CMS with these steps.

The design and development of the IRF PPS reported here were based on a merged file of discharge abstracts from HCFA (the Medicare Provider Analysis and Review [MEDPAR] file) and abstracts containing FIM data from the Uniform Data System for medical rehabilitation (UDSmr), Caredata.com, and HealthSouth. The data describe discharges in calendar years 1996 through 1999 and are further discussed in Section 2. That section also describes our major derived variables, the completeness and representativeness of the data files, and the methods used in our payment simulations.

Payment System Elements

We considered options for the IRF PPS that were based on varying each of the elements of the payment system. Options for the major divisions of the classification system for typical cases are discussed in Section 3. Section 4 discusses subdividing case classes by comorbidities. Section 5 discusses the unusual cases that were not used in developing the main classification system. It examines their costs and discusses options for appropriate payments.

The PPS payment accounts for patient-level variation in need for rehabilitation resources as measured by weights assigned to each CMG. Options for the method used to calculate these weights are discussed in Section 6. Payments are further adjusted based on hospital characteristics that affect costs. Options for hospital adjustment factors are discussed in Section 7.

We used simulations to examine how the payment system elements fit together in a single payment system and to evaluate the likely outcome of the integrated payment system. Simulations were also used to evaluate options for hospital adjustment factors; these simulations are discussed in Section 7. Section 8 describes simulations that update our interim report findings on outlier policy. Section 9 discusses the statistical estimates of hospital case mix used to help calculate the national conversion factor. That section also describes the data used for the impact analyses published in the final rule (CMS, 2001).

Section 10 discusses our future work on monitoring. Future work on refinement of individual payment elements is discussed within Sections 3 through 7.

2. DATA AND METHODS

In the first subsection below, we describe the various data files that we used in our analyses. These files are all updates of those used in our interim report. Our primary data file provides case mix and cost data on Medicare discharges from facilities that were paid under TEFRA as rehabilitation facilities. We also used annual cost reports from the Hospital Cost Report Information System (HCRIS) and a file constructed by the Office of the Actuary that projects the HCRIS data into FY 2002. We constructed a file of hospital characteristics that we used to analyze hospital costs and the likely outcome of policy alternatives.

We also describe Medicare bills for other services provided to beneficiaries. We used records for acute care inpatient discharges to develop the estimates of each hospital's CMI, as discussed in Section 9. As part of the monitoring plan described in Section 10, we intend to use inpatient bills as well as bills for all forms of post-acute care.

After we discuss our data sources, we define a series of variables that are used throughout the analysis and discuss the frequency of missing data. In the third subsection, we describe the sample sizes available for particular analyses and the representativeness of these samples. We conclude Section 2 with a description of the methodology for our simulations. Methods used in individual sections are discussed within those sections.

DATA SOURCES

Case Mix Data

The information in our case mix file comes from discharge abstracts collected on all Medicare patients by HCFA in the course of administering the program and from additional patient information, including the Functional Independence Measure (FIM), that was recorded by a subset of rehabilitation hospitals. The case mix file is described in more detail by Relles and Carter (forthcoming).

Claim Records. HCFA sent us records of all discharges for calendar years (CYs) 1996 through 1999 from the Medicare Provider Analysis and

Review (MEDPAR) file. These records provide patient characteristics, admission and discharge dates, and charges for services rendered during the stay.

We used the provider number in the range 3025-3099 to identify discharges from freestanding rehabilitation hospitals, and the provider code "T" to identify discharges from rehabilitation units of acute care hospitals that were exempt from the acute care PPS. These discharges form the universe of rehabilitation cases that we wished to study.

The size of this universe is shown in Table 2.1. There were 390,048 rehabilitation discharges in CY 1999. Freestanding hospitals accounted for 33 percent of the discharges in each year but constituted only 17.5 percent of the hospital universe in CY 1999.

Table 2.1

Number of Rehabilitation Discharges and Facilities

Calendar Year	Number of Cases			Number of Facilities		
	Total	Free-standing	Excluded Units	Total	Free-standing	Excluded Units
1996	344,126	114,933	229,193	1,081	204	877
1997	359,032	118,541	240,491	1,123	212	911
1998	370,352	122,337	248,015	1,155	214	941
1999	390,048	129,303	260,745	1,165	204	961

Functional Independence Measure and Related Data. Our FIM data come from UDSmr, the Clinical Outcomes Systems (COS) data for medical rehabilitation, and HealthSouth. UDSmr is operated within the Center for Functional Assessment Research, U.B. Foundation Activities, Inc., in Buffalo, New York. COS was operated by Caredata.com, Inc., a provider of health care data and decision support systems, located in Atlanta, Georgia.[1] HealthSouth Corporation is a national provider of health services, with offices in Birmingham, Alabama. COS provided us with its data for CY 1996 and 1997; in early 1998, it ceased to operate.

---

[1] Participation in either UDSmr or COS is/was an entirely voluntary decision by hospital management. Hospitals that were not participating in either database might have used a different version of the FIM, a different assessment instrument, or no assessment process at all.

HealthSouth, however, was one of COS's major clients and provided us data for CY 1998 and CY 1999 from the HealthSouth corporate database, thus providing us with complete 1996-1999 coverage for many of its hospitals.

The data found in all three FIM databases include demographic descriptions of the patient (birth date, gender, zip code, ethnicity, marital status, living setting), clinical descriptions of the patient (condition requiring rehabilitation, ICD-9-CM diagnoses, FIM at admission and discharge), and descriptions of the hospitalization (encrypted hospital identifier, admission date, discharge date, payment source, and discharge destination).

Linking Process. The MEDPAR and FIM files that describe the same discharge were linked using a probability matching algorithm developed during an earlier project and described in Relles and Carter (forthcoming). The linking process proceeded in two steps. The first step determined the Medicare provider number(s) corresponding to each facility code in the FIM database. The second step matched FIM and MEDPAR patients within paired facilities using a probabilistic match algorithm. In addition to hospital identity, the variables we used were admission date, discharge date, zip code, age at admission, sex, and race. All these variables are in each of the files, although sometimes in a slightly recoded form (e.g., birth date rather than age).

During the linking process, we dealt with a number of complications. Some of the records in our FIM database were not paid by Medicare. Further, some of the facilities included in our FIM database were not paid as rehabilitation hospitals under TEFRA (they were either SNFs or long-term care hospitals), so even their Medicare patients' records did not belong in our database. In addition, the FIM data do not provide a complete record of activity at these facilities during the sample years. Because many hospitals were joining UDSmr, COS, or HealthSouth during those years, their records may have been present only part of the year, or their data may not have been properly organized for inclusion in the data systems until somewhat later. Also, a few hospitals were in more than one data system at the same time, and we needed to eliminate duplicate records of the same stay.

Match Quality. We judged the quality of the match, compared to what was possible given our data, in two ways. First, we looked at MEDPAR records for providers that appeared in a FIM database throughout each year and calculated the fraction of the MEDPAR records that we were able to match to a FIM record. As shown in the first column of Table 2.2, we were able to match 87.6 percent of such MEDPAR records in CY 1996 and 90.1 percent in CY 1999. Second, we looked at the percentage of FIM records for which Medicare is listed as the primary payer that we were able to match. In CY 1996 we matched 95.7 percent of such FIM records, and in CY 1999 we matched 95.9 percent. Using both measures, the match rate was very similar for each FIM source.

Table 2.2

Match Rates

Year	% of MEDPAR That Matched	% of FIM That Matched
1996	87.6	95.7
1997	89.4	95.4
1998	89.0	95.4
1999	90.1	95.9

Almost exactly half of all CY 1996 MEDPAR rehabilitation discharges and 66 percent of all CY 1999 MEDPAR rehabilitation discharges were matched. We will describe the analysis sample in more detail following discussion of other data sources.

HCRIS Data

Hospital cost reports are contained in HCRIS. The cost report files contain information on costs and charges by cost center, facility characteristics, and utilization. Each record covers a hospital fiscal year, and all the hospital fiscal year records that began during a specific federal fiscal year are kept together in the same file. The file is named according to the number of years since the beginning of the acute care Medicare Prospective Payment System (PPS).

In the analyses reported here, we used the latest files available in December 1999 for PPS 13 (i.e., hospital fiscal years that began during FY 1996 from 10/1/95 through 9/30/96) and the latest files available in July

2000 for PPS 14 and PPS 15 (hospital fiscal years that began during FY 1998). We also used an older version of the PPS 12 Exempt Hospital and Excluded Units Data Set, which was derived from the HCRIS. Thus, the cost report data coincide with the MEDPAR file except for the last year of data. The portion of the discharge file missing a corresponding cost report varies by hospital, typically ranging from 15 months for hospitals whose cost reporting period begins October 1 to six months for hospitals whose cost reporting period begins July 1.

OACT File

This provider-level file contains estimates by the OACT of TEFRA payments and costs during federal FY 2002. The file contains estimates of total capital costs, operating costs, capital TEFRA payments, operating TEFRA payments, and number of discharges for each provider on the file. The file was received in May 2001 and reflects the wage- adjusted cap on the TEFRA limit that was part of the BBRA. It also reflects revised target rates for hospitals that rebased in their fiscal year that began in FY 1998.

Provider File

This file contains one record for any provider that was on the OACT file or had a rehabilitation discharge paid under TEFRA at any time from CY 1996 through CY 1999. There are 1,442 provider numbers on this file.

For each rehabilitation facility, we assembled information from HCRIS files and other HCFA sources that we discuss in more detail in Section 7. The information includes
- a hospital wage index for the area containing the rehabilitation facility
- data on the number of residents assigned to rehabilitation facilities
- data on the percentage of Medicare cases at each rehabilitation facility that are SSI recipients (person-level data were not gathered)
- FY 2002 TEFRA payments and cost from the OACT file just discussed.

Other Utilization Data

We received the entire MEDPAR file, which includes all hospital discharges. We selected the records for rehabilitation hospitalizations from this file, as discussed above, under the heading "Case Mix Data." Then we linked each rehabilitation hospitalization to the immediately preceding acute care hospitalization and kept only acute care stays where the patient was discharged within 30 days of admission. Approximately 95 percent of these acute care stays were discharged to rehabilitation within a day of admission. These acute care stays were a major source of information for our estimate of the case mix index for hospitals and discharges that were not in our FIM database (Section 9). They also provide an important means of monitoring the coding of impairment group during PPS operation.

We also received a wide variety of Medicare bills for each beneficiary who was hospitalized in a rehabilitation facility during 1996 and 1997. These included HHA bills, SNF bills, bills for durable medical equipment, and institutional outpatient claims.

We used these bills to examine the accuracy of the discharge destination field on the MEDPAR and FIM records. We also used the bills to begin analyses of post-inpatient rehabilitation care, which is a prerequisite for the design of the monitoring system, as discussed in Section 10.

VARIABLE DEFINITIONS AND DATA QUALITY

Variables Used for Classification

The payment amount for each case in the fully phased-in IRF PPS will be determined in large part by the CMG to which it is assigned. For typical cases--i.e., non-transfer, non-death cases--the CMGs that we have studied are a variant of the Function Related Groups with Comorbidities (FRGCs). The first partition in creating FRGCs is the Rehabilitation Impairment Category or RIC, a grouping of codes that describe the impairment that is the primary cause of the rehabilitation hospitalization.

The codes for the primary impairment are identical in the UDSmr and HealthSouth data. They are very similar, but not identical, in the COS

data sets. Most of the differences relate to distinctions within RICs and thus did not affect our analysis. For example, in 1997, UDSmr used codes 8.11 for Status Post Unilateral Hip Fracture and 8.12 for Status Post Bilateral Hip Fracture, while COS used only 8.1 for Status Post Hip Fracture. Although most of the coding differences did not affect our analyses, UDSmr added several new impairment codes related to medically complex conditions in July 1997, while COS used the same codes with an entirely different meaning--to denote cases with multiple impairments. Fewer than 1 percent of COS cases had these codes. We grouped all these cases into the "miscellaneous" RIC because there were not enough cases to analyze separately. In addition, there were 328 COS records with impairment codes that could not be assigned to a RIC in 1996 (1.26 percent of all COS records in 1996) and 323 such cases in 1997 (0.84 percent of COS records).

RICs were created based on clinical criteria and, except for the miscellaneous group, did not put patients who were clinically different from one another in the same RIC. In our 1997 analyses, we used the RICs as defined in version 2 of the FRGs (Stineman et al., 1997b). In our interim report, we evaluated these RICs and updated them to include a separate RIC for burns and a change in the assignment of the multiple fracture codes. The final mapping of impairment codes to RICs is found in Section 3.

Each RIC is subdivided based on age and functional status. Age is taken from the MEDPAR and equals age in years on the day of admission. Functional status is measured by the Functional Independence Measure, or FIM. The FIM is an 18-item measure covering six domains: self-care (six activities of daily living), sphincter control (two items on bowel and bladder management), mobility (three transfer items), locomotion (two items on walking/wheelchair use and stairs), communication (two items on comprehension and expression), and social cognition (three items on social interaction, problem solving, and memory). All 18 items are scored into one of seven levels of function ranging from complete dependence (1) to complete independence (7). In Section 3 we examine the ability of different aggregations of FIM items to predict cost.

There is a slight difference among our data sources in the way three FIM items are recorded. In COS and HealthSouth, separate fields are used to describe independence in walking and in wheelchair use. In UDSmr, a single item covers locomotion in either modality (with a flag saying which is used most frequently or that both are used equally). Similarly, COS and HealthSouth use separate fields for auditory comprehension and visual comprehension and for vocal and nonvocal expression, whereas UDSmr uses one item for comprehension and one for expression. In the COS and HealthSouth data, almost always only one of each of these pairs of items was filled in, so we took whichever one was present. In the cases when both fields for a single FIM item were filled in, we averaged them.

The final partition in our classification is comorbidity. We used ICD-9-CM codes from the MEDPAR and FIM records to analyze the effect of comorbidity. MEDPAR contains up to nine secondary diagnoses on each discharge record to describe comorbidities and complications. UDSmr provides up to seven possible fields, HealthSouth only one, and COS none. Section 4 describes our approach to compensating for the differences among the data sources.

Section 3 describes how we used impairment code, age, and FIM items to develop and analyze alternative classification systems and to develop a recommendation for a classification system to be used in the IRF PPS. Section 4 shows how we used comorbidities to refine this classification system.

Cost per Case

We used the departmental method to estimate the accounting cost of MEDPAR discharges. This method combines MEDPAR information about charges in each ancillary department with the departmental cost-to-charge ratio (CCR) from the cost report to estimate costs incurred by the patient in the department (see, for example, Newhouse, Cretin, and Witsberger, 1989). Separate per diems for routine and special-care days are combined with MEDPAR counts of such days to estimate routine and nursing costs. Special-care days are days spent in intensive care units or in coronary care units. Fewer than 1 percent of rehabilitation days are spent in such units.

The CCRs and per diems were calculated from the PPS 12, 13, 14, and 15 cost reports. The cost report that includes the date of the discharge was chosen for each case if it was available. If it was not available, the cost report closest in time to the date of discharge was chosen. The per diems were inflated (or deflated) from the midpoint of the fiscal year to the day of discharge based on the observed rate of increase in hospital per diems (1.1 percent annually).

Hospitals that charge using an all-inclusive payment rate[2] were omitted from analyses of case-level cost, and their data were not used in calculating weights or payment parameters.

Some departmental CCRs were missing or were found to be outside a plausible range, probably reflecting an error in the cost reporting data. We replaced individual CCRs for all departments except anesthesiology when they were either greater than 10 or less than 0.05. For anesthesiology the CCR was replaced only when greater than 10 or less than 0.01. The replacement was calculated as the mean value of the CCR for the same department within the same type of hospital (either freestanding or unit). There were very few replacements. The largest number occurred in the supplies department: Three replacements of the CCR for supplies occurred in PPS 14 out of 588 hospitals for which we have FIM data; two replacements occurred in PPS 13 out of 585 hospitals for which we have FIM data. In all other departments combined, there were eight replacements in PPS 14 and a similar number in PPS 13.

Routine care per-diem rates were imputed in one sample unit in PPS 14[3] and in a different unit in PPS 15. The first case had missing data for routine cost; its rate was imputed at the average for units in that year. The second unit had a per diem in excess of twice that of the next-closest hospital; its rate was also imputed at the average for units in that year. Special-care units are rarely used for rehabilitation patients, but those rates appeared reasonable for hospitals that have special-care days.

---

[2] In the PPS 15 file, only 45 rehabilitation facilities (3.8 percent) are listed as all-inclusive providers.

[3] This problem did not appear in the smaller cost report file used for the NPRM.

Wage-Adjusted Cost per Case

In creating the final classification system and in all analyses reported in Sections 3 through 6, we used wage-adjusted cost per case. The wage index used in the adjustment is prior to reclassification and reflects the elimination of teaching salaries, as we discuss in Section 7. We used a labor share of 70.5 percent, which matches the time period of the data. Therefore,

$$\text{wage-adj-cost} = \text{cost} / (0.295 + 0.705 * \text{reh\_wi}) \, ,$$

where reh_wi is the wage index for the area containing the hospital.

Discharge Destination

In our interim report, we evaluated alternative ways of determining where the patient went following discharge from inpatient rehabilitation. We examined three potential methods of making this determination. First, the MEDPAR contains a field called "discharge destination," which distinguishes discharges to home, acute care hospitals, SNFs, and other destinations. Second, each of the FIM instruments provides information about the living setting or location to which the patient was discharged, distinguishing many of the same settings as MEDPAR and a few additional ones.[4] Third, we can use Medicare bills to determine which patients went to most of the settings that will classify the patient as a transfer: acute care, other hospital, or SNF. Medicare bills cannot identify transfers to non-SNF nursing homes.

A comparison of the discharge living setting as coded on the FIM instruments with the discharge destination coded on MEDPAR showed that the instruments were recording different outcomes for many cases (see Table 2.2 of our interim report, Carter et al., 2000). We therefore used the Medicare bills for acute hospitalizations and for SNF stays to examine the relative accuracy of FIM and MEDPAR discharge location (Carter et al., 2000, Tables 2.3 and 2.4). The FIM discharge setting corresponded more closely than MEDPAR discharge destination to the information found on the bills.

---

[4] The setting choices on the FIM instruments are similar but not identical. Nevertheless, by grouping the codes on each instrument, one can define comparable settings.

Based on these analyses, we concluded that the FIM discharge
destination is more accurate than the MEDPAR data but that neither is
completely accurate. In addition, we decided to group together discharges
to SNFs with other discharges to long-term care. We used the FIM discharge
settings 6 though 9 to identify transfers to hospitals and settings 4, 5,
12, and 13 to identify transfers to long-term care. Two decisions
remained: What to do with the cases whose discharge setting was 10
("other") in the UDSmr and HealthSouth data, and what to do when the FIM
discharge setting was missing. In creating FRGs, we selected only cases
whose discharge setting was neither 10 nor missing. For other work, we
took the most likely meaning of the data and treated 10 as "discharged
home." When the discharge location on the FIM instrument was missing,[5] we
used the MEDPAR discharge destination information. We also allowed the
Social Security Administration's definition of date of death to override
discharge destination. When the MEDPAR provided a verified date of death
equal to the date of discharge or the preceding day, we counted the
patient as having died in the rehabilitation hospital.

ANALYSIS SAMPLE

Here we describe the completeness of the case mix records and the
cost data, and therefore the sample sizes available for analysis. Almost
all of our analyses were done with the matched FIM/MEDPAR file with good
case cost data. However, the impact analysis and the estimate of a budget
neutral conversion factor were based on all hospitals in the OACT file for
FY 1999 for which CMI could be estimated (see Section 9). The impact
analyses for the NPRM were based on 1997 cases in both the matched
MEDPAR/FIM database and the OACT data. Below we describe how
representative these analysis files are of the entire set of
rehabilitation discharges paid under TEFRA from 1996 through 1999.

---

[5] In the UDSmr, the code for "missing" is 99; in the COS, it is a
blank; in HealthSouth, it is 0 or a blank.

Sample Size

Table 2.3 shows the size of the sample available for analysis. The population is the count of MEDPAR rehabilitation stays paid under TEFRA in each year. We matched FIM data with 49.9 percent of the 1996 records and had some matched records from 52.3 percent of hospitals. The number of hospitals participating in the source FIM databases grew each year--by 1999, we had matched 65.9 percent of cases and had data from 62.1 percent of hospitals.

Several hundred cases annually had incomplete data for either the impairment that was the primary cause of the hospitalization or one of the items that entered into the motor or cognitive score. The remaining cases are counted in the table as being "Good FIM."[6] In each year there were several thousand cases for which we could not estimate a case cost and which could not be used in most of our analyses. Most of these were from all-inclusive providers for whom we lost all cases.[7] The remaining records with good FIM and cost data are shown in Table 2.3 in the first row labeled "Case cost and FIM."

---

[6] In the interim report, we imputed FIM data for cases with some, but not all, FIM items filled in. Thus, the numbers of 1996 and 1997 cases in the analysis files for this report are slightly lower than those used for the interim report.

[7] There were also several cases for which there were no charges and therefore no way to estimate ancillary costs.

Table 2.3

Analysis Sample Size

| | Discharges | | | | | | | | Hospitals | | | | | | | |
| | 1996 | | 1997 | | 1998 | | 1999 | | 1996 | | 1997 | | 1998 | | 1999 | |
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	N	%
Population	344,126	100.0	359,032	100.0	370,352	100.0	390,048	100.0	1,081	100.0	1,123	100.0	1,155	100.0	1165	100.0
Matched records	171,626	49.9	206,032	57.4	232,691	62.8	257,024	65.9	565	52.3	631	56.1	686	59.4	724	62.1
Good FIM	170,811	49.6	205,092	57.1	232,310	62.7	256,702	65.8	565	52.3	631	56.1	686	59.4	724	62.1
Case cost and FIM	167,081	48.6	201,372	56.1	228,247	61.6	249,941	64.1	556	51.4	619	55.1	668	57.8	695	59.7
Excluding Maryland	171,298	49.8	205,542	57.2	232,128	62.6	256,401	65.7	564	52.2	630	56.1	685	59.3	723	62.1
Good FIM	170,483	49.5	204,603	57.0	231,747	62.5	256,079	65.7	564	52.2	630	56.1	685	59.3	723	62.1
Case cost and FIM	166,753	48.5	200,883	56.0	227,684	61.5	249,318	63.9	555	51.3	618	55.0	667	57.7	694	59.6
Impact analysis			172,104				347,809	89.2							1024	87.9

Our classification analyses used data from all available hospitals. However, hospitals in Maryland will not be included in the IRF PPS; consequently, we eliminated data from the single Maryland hospital that was in our data set when we calculated parameters of the IRF PPS such as weights. We call the subset of non-Maryland discharges with good FIM data and case costs our *analysis sample*. Many of our parameter estimations (such as case weights) and our simulations are based on the CY 1999 analysis sample: 249,318 cases from 694 hospitals (63.9 percent of the population of cases, 59.6 percent of the population of hospitals).

For the impact analyses in the NPRM and the interim report, we were restricted to hospitals for which the OACT predicted costs and TEFRA payments and for which we had FIM and cost data. The size of the sample with all required data is shown in the 1997 column of the last row of Table 2.3. For the simulations of FY 2002 payments using 1999 FIM and MEDPAR data, we were able to estimate case mix for many more hospitals (see Section 9). The same impact analysis sample was used as a primary input to the conversion factor calculation. In the following subsection, we examine the representativeness of this impact analysis sample and of our full analysis sample. (The representativeness of the impact analysis sample used in the interim report is described in that report.)

Representativeness of Sample

Table 2.4 shows how well our samples represent the patient characteristics found in the universe of all inpatient rehabilitation cases paid under TEFRA. The analysis sample describes 49 percent of all 1996 rehabilitation discharges and 64 percent of 1999 discharges. Both samples contain a reasonable representation of all demographic groups, although they slightly underestimate minorities, those 95 or older, and the aged with end-stage renal disease (ESRD).

Table 2.4

Percentage of Inpatient Rehabilitation Discharge Universe in Samples,
by Patient Characteristics

Patient Characteristic	1996		1997		1998		1999	
	No.	% in Analysis Sample	No.	% in Analysis Sample	No.	% in Analysis Sample	No.	% in Analysis Sample
Total	344,126	48	359,032	56	370,352	61	390,048	64
Sex								
Male	214,727	48	223,251	56	230,833	61	244,008	64
Female	129,399	49	135,781	56	139,519	62	146,040	64
Race								
White	296,894	49	309,026	56	318,176	62	335,536	64
Black	35,090	46	37,491	53	38,459	60	39,671	62
Other	12,142	48	12,515	52	13,717	55	14,841	59
Age								
<65	27,215	48	29,667	56	31,573	61	33,359	64
65-69	48,413	49	48,570	56	48,759	62	49,192	64
70-74	68,219	49	70,529	56	70,874	62	71,842	65
75-79	78,885	49	82,851	57	84,700	62	89,266	64
80-84	65,863	48	68,854	56	71,371	62	76,094	64
85-89	39,217	47	41,123	55	43,934	60	48,886	63
90-94	13,549	45	14,390	52	15,794	58	17,531	62
95+	2,765	42	3,048	49	3,347	57	3,878	59
Beneficiary status								
Aged without ESRD	312,482	49	325,114	56	334,287	62	352,106	64
Aged with ESRD	3,424	43	3,097	46	3,422	50	3,462	53
Disabled without ESRD	24,836	48	27,176	56	28,835	61	30,465	63
Disabled with ESRD	2,025	50	2,144	57	2,311	63	2,471	65
ESRD only	1,359	46	1,501	55	1,497	60	1,544	63

- 24 -

Tables 2.5 and 2.6 show how well our samples represent the hospital characteristics in the Medicare universe. In Table 2.5, we count sample discharges; in Table 2.6, we count each hospital equally: A hospital is counted in the sample if it had *any* sample discharges in a given year. The differences between sample and universe are somewhat larger in hospital characteristics than in patient characteristics. In particular, our analysis sample underrepresents units of acute care hospitals, rural facilities, and those with a high proportion of low-income patients. On average, freestanding facilities are larger than units. Within each group our sample is drawn slightly more from the larger facilities.

The sample used for the impact analyses and calculation of the conversion factor is much larger and more representative. It includes 89 percent of cases and 88 percent of hospitals. The percentages in the sample are close to equal across both freestanding facilities and units, location type, census region, and teaching status (when known). The conversion factor sample, however, still somewhat underrepresents very small hospitals and units.

SIMULATIONS

We simulated a variety of alternative designs for the IRF PPS. Most of the simulations were based solely on CY 1999 data and were designed to test individual features of an IRF PPS such as options for the amount of outlier payments, or facility adjustments. Each simulation was based on the assumption that changes in the payment method would not affect provider behavior.

The CY 1999 discharges in the merged MEDPAR-FIM analysis file were used in the simulations after selected interrupted stays were bundled as described in Section 6.[8] Each simulation paid exactly the same amount of funds (to within one dollar per case). We chose the conversion factor so that total payments would equal total costs.

[8] We made some policy recommendations based on simulations that used different bundling rules. However, in this report we present the final analyses, which show that the policy recommendations are consistent with our best data and with other policy decisions.

Table 2.5

Percentage of Inpatient Rehabilitation Discharge Universe in Samples,
by Hospital Characteristics

Hospital Characteristic	1996		1997		1998		1999		
	No.	% in Analysis Sample	No.	% in Analysis Sample	No.	% in Analysis Sample	No.	% in Analysis Sample	% in Impact Sample
Total	344,126	48	359,032	56	370,352	61	390,048	64	89
Unit	229,193	45	240,491	50	248,015	54	260,745	55	90
Freestanding	114,933	55	118,541	68	122,337	77	129,303	82	88
Location type									
Large urban	170,351	44	174,293	52	177,938	58	181,642	62	90
Other urban	144,048	56	153,567	64	158,473	69	168,862	70	90
Rural	29,656	34	31,106	40	33,196	45	36,308	48	86
Location									
New England	17,397	33	17,760	54	17,291	69	17,958	68	88
Middle Atlantic	69,654	46	69,658	51	71,043	61	73,242	64	91
South Atlantic	57,893	57	61,876	68	64,440	72	66,968	73	88
East North Central	58,847	60	61,611	63	62,214	65	64,557	66	93
East South Central	25,736	56	27,793	63	28,875	67	30,203	70	89
West North Central	21,436	40	22,206	47	22,957	56	24,107	62	91
West South Central	55,762	39	60,599	46	64,744	57	68,393	59	87
Mountain	15,841	33	15,860	48	16,577	45	18,680	54	84
Pacific	21,560	51	21,669	52	21,486	53	23,024	55	95
Low-income percentage									
<10	214,202	49	222,668	56	228,387	60	243,530	62	87
10-19	99,403	50	104,042	59	107,956	66	112,036	70	94
20-29	20,938	45	22,475	52	23,647	60	24,560	59	87
=>30	9,583	27	9,847	35	10,362	43	9,922	52	92

Table 2.5 (cont.)

Hospital Characteristic	1996		1997		1998		1999		
	No.	% in Analysis Sample	No.	% in Analysis Sample	No.	% in Analysis Sample	No.	% in Analysis Sample	% in Impact Sample
Average daily census (ADC)									
Freestanding									
<25	9,511	48	9,925	52	10,744	57	11,516	57	73
25-49	36,335	51	39,282	63	41,678	80	45,008	86	92
=>50	69,087	58	69,334	73	69,915	79	72,779	83	88
Unit									
<10	42,403	35	43,051	38	44,099	42	51,503	42	76
10-24	116,003	43	124,915	49	127,024	51	131,334	52	94
=>25	70,787	54	72,525	58	76,892	65	77,908	70	92
Teaching: interns and residents to ADC									
None	281,019	49	295,634	56	302,821	62	316,595	65	90
<0.05	30,084	51	30,961	64	31,206	64	32,080	73	98
0.05-0.15	20,230	43	21,030	45	21,509	53	21,410	56	89
=>0.15	11,124	43	10,868	50	13,646	59	14,201	59	87

Table 2.6

**Percentage of Inpatient Rehabilitation Hospital Universe in Samples,
by Hospital Characteristics**

Hospital Characteristic	1996		1997		1998		1999		
	No.	% in Analysis Sample	No.	% in Analysis Sample	No.	% in Analysis Sample	No.	% in Analysis Sample	% in Impact Sample
Total	1,081	51	1,123	55	1,155	58%	1,165	60	88
Unit	877	49	911	53	941	53%	961	55	89
Freestanding	204	62	212	65	214	78	204	81	82
Location type									
Large urban	515	49	538	53	545	55	535	60	91
Other urban	416	60	431	64	443	67	437	68	90
Rural	149	35	152	40	157	46	164	48	87
Location									
New England	38	53	40	58	42	64	40	65	90
Middle Atlantic	149	38	156	47	156	50	159	57	90
South Atlantic	144	63	148	66	146	72	144	74	92
East North Central	209	64	221	65	218	67	217	68	92
East South Central	54	65	54	67	56	66	57	65	89
West North Central	96	44	100	50	102	56	100	60	92
West South Central	199	40	210	46	225	51	222	53	84
Mountain	72	39	73	42	78	42	75	51	87
Pacific	120	57	121	55	123	56	125	57	95
Low-income percentage									
<10	656	51	677	55	701	57	724	57	84
10-19	296	57	309	60	312	64	308	69	95
20-29	77	45	82	50	87	55	82	56	90
=>30	52	31	55	31	55	35	51	37	98

- 28 -

Table 2.6 (cont.)

Hospital characteristic	1996		1997		1998		1999		
	No.	% in Analysis Sample	No.	% in Analysis Sample	No.	% in Analysis Sample	No.	% in Analysis Sample	% in Impact Sample
Average daily census (ADC)									
Freestanding									
<25	62	37	64	41	67	46	59	49	61
25-49	73	68	78	69	78	94	76	96	93
=>50	69	78	70	83	69	91	69	91	88
Unit	307	39	321	40	346	41	371	44	78
<10									
10-24	432	51	450	57	455	57	452	58	96
=>25	138	63	140	67	140	71	138	75	95
Teaching: interns and residents to ADC									
None	912	52	961	54	979	58	965	61	90
<0.05	53	68	55	80	57	72	57	75	98
0.05-0.15	48	52	52	54	52	58	51	59	100
=>0.15	46	46	47	53	49	55	49	57	92

We evaluated each simulation with respect to an outcome at the hospital level. For the interim report, we also evaluated case-level data. However, as we explain in Section 6, there are reasons to think that our case cost data are somewhat compressed. Thus we believe that case-level accuracy is a less important criterion for judging payment options than are hospital-level data.

Payment for Groups of Hospitals

We calculated the payment-to-cost ratio using a simulated payment and the estimated cost of the case using the departmental method described above. The aim of the IRF PPS is to pay each rehabilitation facility in proportion to the costs of efficiently producing the care required by its set of Medicare patients. We would not want to pay hospitals for inefficiency or even for a greater intensity of care than is typically received by patients with similar clinical characteristics and social support levels. We follow Gianfrancesco (1990) and others in using the relationship between reimbursement and costs for groups of hospitals as one way of detecting systematic unfairness in a payment system. The relationship between reimbursement and costs for a group is summarized by the ratio of average simulated payment to the average estimated cost of the case using the departmental method described above. A finding of a higher ratio of payment to cost in one group than in another would suggest unfairness in the payment mechanism, unless there was a good reason to believe that efficiency differed between the groups.

We present payment-to-cost ratios and a comparison of PPS payments and TEFRA payments for groups of hospitals defined by

- rural, urban, and large urban location
- census region
- freestanding versus exempt unit
- size of the low income population
- teaching status
- size of Medicare rehabilitation program
- age of facility.

We summed payments and costs for all cases in the group in order to calculate the ratio. Thus all ratios are "case weighted" rather than "hospital weighted."

Accuracy at the Hospital Level

The accuracy with which the IRF PPS matches payment to cost at the hospital level was measured by the mean of the absolute value of the difference between each hospital's average cost per case and its average payment per case. We also report the root mean square error (RMSE) of payment as a prediction of cost. This measure gives greater weight to larger errors. The Efron R-squared is 1 minus the ratio of MSE to the variance of average cost per case.

Financial Risk

Rehabilitation hospitals are typically much smaller than acute care hospitals. The average freestanding rehabilitation hospital discharges only 634 Medicare cases per year, and the average exempt unit discharges only 271 Medicare cases per year.[9] The median acute care PPS hospital discharged over 1,200 patients in the same timeframe. The smaller size of rehabilitation facilities suggests that they would be at higher financial risk from a PPS than acute hospitals would, i.e., there would be a higher likelihood that their costs would exceed revenues by a substantial amount. However, the higher risk from their small size is partially offset by the greater homogeneity of the cost of cases within FRGCs than within DRGs. Our previous study (Carter, Buchanan, et al., 1997, Table 5.6) showed that, in the absence of outlier payments, the financial risk faced by the typical freestanding rehabilitation hospital under an acute IRF PPS would be quite similar to the amount of risk faced by the typical acute care PPS hospital in the absence of outlier payments. Because exempt units are quite a bit smaller than freestanding rehabilitation hospitals, their risk is larger--close to that of rural hospitals under the acute care PPS. In the sample used in our interim report, average risk in the absence of outlier payments was slightly smaller than in the 1997 study (3.03 versus 3.19).

To assess the extent of financial risk from an IRF PPS, we used the following measure of financial risk under a PPS for a particular hospital. A hospital's risk is *the standard deviation of annual profit around its expected value expressed as a percentage of annual revenues* (Keeler,

---

[9] Calculated from the calendar year 1999 MEDPAR file. Medians are slightly lower.

Carter, and Trude, 1988). Profit is defined as *Medicare revenues minus Medicare costs*. We estimated this quantity by assuming that each hospital has its own population of cases that might appear for admission and that actual cases are drawn independently from this distribution. Let

s_i = an estimate of the standard deviation of profit per case at hospital i,

n_i = number of annual cases at hospital i, and

r_i = average revenue per case at hospital i.

Then, under our assumption of a random draw of cases,

$$Risk_i = s_i / (sqrt(n_i)r_i).$$

Actual year-to-year variation in profits in the acute care PPS has been shown to be somewhat higher than the variation estimated from this model (Carter and Farley, 1993). At least in part, this was due to management actions that affected costs or revenues.

This risk measure has several characteristics that make sense. Risk decreases as hospital size increases for hospitals with a similar case mix. And the greater the variability of profit among cases, the greater the risk, since variability increases the chance that the hospital will receive so many unprofitable cases that it cannot offset its losses with gains.

In implementing the risk measure, we used the simulated payment data on matched MEDPAR-FIM cases in the analysis sample to estimate s_i and r_i, but used all discharges in the MEDPAR file to estimate n_i.

Impact Analyses

We also simulated hospital-level payments to examine the impact of the IRF PPS relative to projected costs and projected TEFRA payments in FY 2002. The method of determining the conversion factor is discussed in Section 9. We measured the impact by examining the ratio of payment to cost and the ratio of PPS payment to TEFRA payment for groups of hospitals. We used the OACT estimate of FY 2002 cost per case and TEFRA payment per case.

3. CASE CLASSIFICATION SYSTEM

INTRODUCTION

Case classification is a major step in developing a PPS payment formula. In this section, we describe the construction of a set of Functional Independence Measure-Function Related Groups (FIM-FRGs, or simply FRGs). FRGs partition the population into groups that are medically similar and that have similar expected resource needs.

Facilities will ultimately be compensated for typical cases (i.e., cases that are discharged to the community after a full course of rehabilitation) according to a formula that depends primarily on their assigned class, adjusted by comorbidities, area wage rates, and other hospital characteristics. Here, we classify only typical cases, which we define more precisely in Section 6. (In Section 5, we discuss payment rules for unusual cases, including interrupted stays.) The discharge is our unit of classification: For typical cases, a case and a discharge are the same.

We built on the FRG classification methodology developed in Stineman et al. (1994), extended in Carter et al. (1997), and further extended in Carter et al. (2000). Each source applies classification and regression trees (CART) to develop a set of FRGs tailored to the latest data and incorporating new refinements. Our 1997 report developed FRGs from 1994 data using the 20 rehabilitation impairment categories (RICs) proposed by Stineman. It also verified that the original Stineman FRGs were stable, effective predictors of resource use. The 2000 report (referred to as the "interim" report) refit FRGs from 1996 and 1997 data, confirmed the predictive power of FRGs on other years' data, and explored the utility of reformulating the RICs. This section uses four years of data (1996 through 1999), reconfirms the predictive power of the CART-based FRGs, and evaluates the quality of FRGs based on their performance relative to a "gold standard" model.

Our final classification system is based on the structure of the FRGs enhanced with information about comorbidities (Section 4). FRG development precedes inclusion of comorbidity information and requires two steps:

1. *Grouping cases that are clinically similar.* Here we started by using the 21 RICs developed in the interim report, Stineman's 20 RICs, plus another RIC for burns. We recommended adding the burn RIC after we had examined a

variety of changes that Dr. Stineman suggested might improve either clinical or resource homogeneity.

2. *Grouping cases that have similar resource needs.* Within RICs, we used the statistical method CART (Breiman et al., 1984) to partition the population of cases into groups that were homogeneous with respect to resource use and functional impairment.

We begin with a brief review of the classification results from the interim report. Then we provide a description of the data, a subsection on modeling methods and results, and finally a subsection on obtaining new FRGs.

REVIEW OF PREVIOUS CLASSIFICATION SYSTEMS

In the project's interim report, we developed a set of FRGs based on 1996 and 1997 data. These FRGs were meant to support development of the payment formula used in the NPRM and to invite feedback and criticism. We intended to update them based on comments received and on the arrival of 1998 and 1999 data.

Developing this classification system entailed experimenting with alternative ways to define RICs and developing FRGs as predictive models of resource use within the newly defined RICs. We briefly review our findings below.

Rehabilitation Impairment Categories

FIM data contain an "impairment" code that gives the reason for the rehabilitation stay. Stineman et al. (1994) mapped these codes into rehabilitation impairment categories (RICs). The 1997 study by Carter et al. convened a panel of rehabilitation experts, who generally approved the RIC definitions but offered some suggestions to try if sample sizes became larger. Further, Dr. Stineman wished to explore some modifications to the definition of RIC in a larger sample.

These partitions were reconsidered in the interim report. Because more data were available than had been when the earlier set of RICs was generated, we felt that we should examine different alternatives. We redefined RICs based on clinician judgment of the clinical homogeneity of the patients, backed up by analyses of resource costs.

We then examined new RICs that split, combined, or rearranged existing RIC groupings. The criterion for whether an additional grouping would be desirable was whether it would lead to more accurate predictions. We evaluated this by fitting cost prediction models within alternative candidate RICs, then seeing if

- 35 -

there was improvement in the mean cost predictions accompanied by a drop in root mean square error. In most cases, we saw very little change in performance; so in view of the already broad acceptance in the rehabilitation community of the existing RICs, we chose to leave those RICs alone. There were also several cases where the new groupings performed substantially worse than the older grouping.

Of the twelve alternative groupings we tried, there were two areas where positive changes seemed large enough to be important, and one case in which a cosmetic change seemed warranted:

1. We defined a new RIC for burns and eliminated burn cases from the "miscellaneous" RIC.

2. We moved the group "status post major multiple fractures" from the "orthopedic--lower extremity fracture" RIC to group it with other cases in the "major multiple trauma, no brain or spinal cord injury" RIC.

3. Switching RIC 20 (major multiple trauma with injury to brain or spinal cord) and RIC 18 (miscellaneous) enabled us to put the major multiple trauma RICs together in numerical sequence.

Our final definition of impairment codes is shown in Table 3.1.

Function Related Groups (FRGs)

We used the statistical method CART, as described in Breiman et al. (1984), to partition the population of cases within RICs into groups that appeared homogeneous with respect to resource use and functional impairment. CART was also used in Stineman's initial development of the FRGs and in our 1997 evaluation of the Stineman methods. The CART algorithm examines a set of independent variables and searches for a partition that explains variation in the dependent variable. The algorithm is recursive. At each step, CART examines all possible two-way splits of the existing groups. It chooses the split that offers the greatest increase in R-squared for that step. CART stops splitting when it thinks it is introducing just noise.

Table 3.1

Final Grouping of Impairment Group Codes into Rehabilitation Impairment Categories

Rehabilitation Impairment Category		Impairment Groups
1	Stroke	1.1 through 1.9
2	Traumatic brain injury	2.2, 2.21, 2.22
3	Nontraumatic brain injury	2.1, 2.9
4	Traumatic spinal cord	4.2, 4.21 through 4.23
5	Nontraumatic spinal cord	4.1, 4.11 through 4.13
6	Neurological	3.1, 3.2, 3.3, 3.5, 3.8, 3.9
7	Hip fracture	8.11 through 8.3
8	Replacement of lower extremity joint	8.51 through 8.72
9	Other orthopedic	8.9
10	Amputation, lower extremity	5.3 through 5.7
11	Amputation, other	5.1, 5.2, 5.9
12	Osteoarthritis	6.2
13	Rheumatoid, other arthritis	6.1, 6.9
14	Cardiac	9
15	Pulmonary	10.1, 10.9
16	Pain syndrome	7.1 through 7.9
17	Major multiple trauma, no brain or spinal cord injury	8.4, 14.9
18	Major multiple trauma, with brain or spinal cord injury	14.1, 14.2, 14.3
19	Guillain-Barré syndrome	3.4
20	Miscellaneous	12.1, 12.9, 13, 15, 16, 17 through 17.9
21	Burns	11

In the interim report, we used CART's "tenfold cross-validation method" to determine the optimal number of splits in the final classification tree. This method divides the data into ten mutually exclusive sets of equal size, chosen at random. For each set, a tree with k nodes is fit on the other 90 percent of the data, and the square error of the predictions for the other 10 percent is computed and summed over the ten sets. CART then chooses the k with the minimum sum of squares error (equivalently, the maximum R-squared) and fits a tree on the entire data set with k nodes.

Because this method resulted in many splits in large RICs, we traced the values of R-squared as we increased the number of nodes within each RIC. These traces show that the gain in R-squared per node is rather low for those trees exceeding 10 splits. Furthermore, as the size of the tree approaches 20 nodes, the R-squared is often very close to the CART maximum. We used this fact to justify restricting our consideration to models with fewer than 20 nodes in all cases.

Using all our 1997 data, we fit CART to data sets that were at times very large. We required a minimum of 100 cases in each FRG. Raw CART fits produced a total of 359 nodes. This is not too surprising--CART adds nodes as long as the increase in R-squared seems statistically significant. With large samples, even minor differences are statistically significant.

For administrative simplicity, we did not want to create such a large number of groups. In addition, we did not want to create groups characterized by very small intervals of motor or cognitive scales for fear it would encourage upcoding. Because it would not be enough to simply use what CART considered the "best" model, our strategy was to produce some reduced-size models according to their perceived statistical power and practical importance. To accomplish this, we employed four steps:

1. We looked at the R-squared trace produced by CART, which confirmed that there is little to gain in going beyond 20 nodes per RIC. This yielded 232 nodes in total. We therefore bounded the number of nodes at 20 and considered further reductions with respect to these bounded models.

2. We used a stopping rule based on one prediction standard error (a rule recommended by Breiman et al., 1984). This reduced the total number of nodes to 143.

3. We tried stopping when R-squared was within .01 of the R-squared for the one-standard-error models. For computational reasons, we used the actual R-squareds, not the cross-validated ones shown in the table. Results were only minimally affected. This reduced the total number of nodes to 104.

4. We looked at the FRG category definitions from the 104-node model and noticed that predicted costs were sometimes quite close (within $1,500) among lower branches of the same part of the tree. We combined these categories, thereby reducing our tree size to 92 nodes.

We examined the progressive reduction in the number of nodes as a result of each of these steps. We selected two candidate models for further evaluation: a 143-node model arising from CART, and a 92-node reduced model arising from requiring little gain in the R-squared trace and the pairing/tripling up of adjacent branches of the classification tree wherever predicted costs were close. We used simulated payments to assess prediction bias for various combinations of demographic and hospital factors. We found that the 143-node model had only a very small effect on overall accuracy and no noticeable effect on payment for any group of hospitals. Thus, we recommended the 92-node model to HCFA, and we went forward with the 92-node model to develop other aspects of our

proposed payment plan (e.g., wage adjustments, outlier payments). The definitions of the groups can be found either in our interim report or in the NPRM.

DATA

For the present study, we used the merged MEDPAR/FIM discharge file for calendar years 1996 through 1999, discussed in Section 2. Construction of the file itself is described in Relles and Carter (2002).

Data Set Contents

The merged MEDPAR/FIM data contained several variables we needed for modeling and classification. Table 3.2 identifies these variables and indicates at which stages of the process they were used.

The selection variables define what we think of as the typical case. We exclude transfers to hospitals and nursing homes, deaths, cases of three days or less duration, and statistical outliers. Also, the clinical partitioning and resource use variables needed to be present and in range. Selection was based on the intersection of the rules shown in Table 3.3.

Table 3.4 shows the amount of data we had to work with, before and after selection. Most of the reduction in cases is for ineligibility: deaths, interrupted stays, or transfers. The last column indicates how many cases were kept with full information. Overall, the reductions due to missing cost data and data quality (present and in-range, excluding statistical cost outliers) are small: about 3 percent in 1996, 4 percent in 1997, 2 percent in 1998, and 3 percent in 1999. Fortunately, the additional reduction due to outliers is especially small, less than 0.3 percent everywhere, so we do not believe we are contaminating our results by the outlier exclusions.

Table 3.2

MEDPAR/FIM Variables and Stages of Use

Purpose	Variable	Source	Description
Selection			
	AGE	MEDPAR	Age
	DISSTAY	FIM	Discharge stay indicator
	LOS	MEDPAR	Length of stay
	IMPCD	FIM	Rehabilitation impairment codes
	PROVCODE	MEDPAR	Provider code
	PROVNO	MEDPAR	Provider number
	TCOST	MEDPAR	Total cost estimates, based on cost-to-charge ratios, adjusted by area wage index[a]
Clinical partitioning			
	IMPCD	FIM	Impairment code
	RIC	FIM	Clinical groupings resulting from impairment code mappings
Resource use			
	TCOST	MEDPAR	Total cost estimates, based on cost-to-charge ratios, adjusted by area wage index[a]
	COGNITIVE	FIM	Cognitive scores[b]
			Comprehension
			Expression
			Social interaction
			Problem solving
			Memory
	MOTOR	FIM	Motor scores[b]
			Eating
			Grooming
			Bathing
			Dressing--upper body
			Dressing--lower body
			Toileting
			Bladder management
			Bowel management
			Bed, chair, wheelchair transfer
			Toilet transfer
			Tub or shower transfer
			Walking or wheelchair
			Stair ascending and descending
	AGE	MEDPAR	Age

[a] These methods are described in Section 2.

[b] These individual components are organized into various types of indices, according to body areas and types of impairment. See Table 3.6.

Table 3.3

Rules for Selection of Modeling Cases

Variable	Selection Requirement
AGE	Between 16 and 105
DISSTAY	Indicates discharged to the community
LOS	More than three days, less than one year.
IMPCD, TCOST	We excluded cases with wage-adjusted log-cost more than three standard deviations from its average within RIC
IMPCD	Contained in impairment list for assignment to rehabilitation categories (see Table 3.1)
TCOST, COGNITIVE, MOTOR	Greater than zero

Table 3.4

Number of Linked MEDPAR/FIM Records

Calendar Year	Matched Records	Present and in-Range	Eligible	Excluding Outliers
1996	171,626	166,645	126,900	126,581
1997	206,032	197,076	148,526	148,142
1998	232,691	228,248	170,266	169,816
1999	257,024	249,941	187,257	186,766

Our numbers of 1996 and 1997 cases are slightly reduced from the numbers shown in Table 3.2 of the interim report, for two reasons. First, in 1996 and 1997 we were working only with the standard motor and cognitive indices and had imputed their values from partial information, if available. Here, because we needed to work with individual components and several alternative subscales, we eliminated all cases that were not complete on all components. This reduced our 1996 counts by about 300 cases and our 1997 counts by about 200. Second, we had allowed discharges to some subgroups of nursing homes in our 1996 and 1997 models. HCFA subsequently decided to classify such cases as transfers, so we adjusted our 1996 and 1997 data sets to exclude them, subtracting about 800 cases in each year.

Case Stratification and Sample Sizes

Previous work had established 21 clinical groupings of patients according to rehabilitation impairment codes within which we would be fitting models. Table 3.5 describes those groupings and the sample sizes available for the modeling effort according to the selection rules in Table 3.3.

Table 3.5

RIC Definitions and Sample Sizes

Rehabilitation Impairment Category	1996	1997	1998	1999
1 Stroke	32,687	35,026	37,012	37,340
2 Traumatic brain injury	1,383	1,629	1,871	2,053
3 Nontraumatic brain injury	2,517	2,863	3,402	3,758
4 Traumatic spinal cord	738	810	930	953
5 Nontraumatic spinal cord	3,782	4,340	5,295	5,837
6 Neurological	4,730	5,717	7,832	8,875
7 Hip fracture	16,017	17,167	18,774	20,627
8 Replacement of lower extremity joint	31,151	37,383	40,931	43,427
9 Other orthopedic	5,292	6,547	8,022	9,310
10 Amputation, lower extremity	4,810	5,423	5,930	6,156
11 Amputation, other	354	477	542	662
12 Osteoarthritis	2,340	2,854	3,983	5,036
13 Rheumatoid, other arthritis	1,169	1,521	1,944	2,350
14 Cardiac	4,097	5,662	6,885	8,104
15 Pulmonary	2,442	3,561	4,340	5,382
16 Pain syndrome	1,321	1,873	2,529	2,993
17 Major multiple trauma (MMT), no brain or spinal cord injury	1,188	1,288	1,540	1,679
18 MMT, with brain or spinal cord injury	156	222	221	256
19 Guillain-Barré syndrome	240	278	299	313
20 Miscellaneous	10,097	13,398	17,423	21,553
21 Burns	70	103	111	102
Total	126,581	148,142	169,816	186,766

MODELING METHODS AND RESULTS

A meeting of the project's technical expert panel was convened in May 2000 to review a draft of our interim report. During this meeting we discussed the methods and results of our initial FRG fits. We took from that meeting a set of three basic suggestions for improving on what we had done:

- *Explore alternative model forms.* Develop models to compete with CART in terms of having strong predictive performance.

- *Consider indices of function in addition to the cognitive and motor scores.* Payment formulas based on these measures might offer better estimates of cost.

- *Evaluate out-of-sample performance of the models.* An important element of a payment system is whether payment formulas offer accurate prospective estimates of cost.

This subsection discusses the methods we used to implement these techniques and discusses our results in sifting through these methods and reaching conclusions about which ones to use in developing a set of recommended FRGs. The description and evaluation of our recommended FRGs are deferred to a later subsection.

Suggestions of the Technical Expert Panel

Explore Alternative Model Forms

We expected that classification and regression trees would form the final determination of the FRGs. According to the Balanced Budget Act of 1997, the rehabilitation PPS system is to be based on discharges classified according to FRGs based on impairment, age, comorbidities, and functional capability of the patient, as well as other factors deemed appropriate to improve the explanatory power of FRGs. CART is the traditional method of generating FRGs (Stineman et al., 1997b) and a reasonable method of determining rules to classify patients into groups that explain cost. Various algorithms have been proposed to build tree-structured regression models, all of which tend to be minor variations on CART. CART is efficient at producing simple and effective rules for prediction but also has its limitations. We discuss CART's strengths and limitations in the next subsection.

Even after computing an unbiased estimate of the predictive performance of a particular regression tree, we found it difficult to judge how much better we might have done if we were not subject to CART's limitations. We know that R-squared ought to be between 0.0 and 1.0 with the highest values indicative of nearly perfect prediction. But when R-squared is potentially much lower than 1.0, we need a way to judge whether CART has performed as best as could be expected. To further investigate this, we compared CART's performance with that of other methods.

We compared CART to ordinary linear least squares regression models, generalized additive models (GAM), and multiple adaptive regression trees (MART). The first of these three methods is classic, the second is relatively new, and the last is the latest in prediction methodology. All these models are discussed in the statistical literature. We used the version of GAM (Hastie and Tibshirani, 1990) implemented in the statistical package S-plus. MART is described in Hastie, Tibshirani, and Friedman (2001); we used software developed by one of its co-authors.

We assessed each model's predictive performance on preceding and subsequent years. That is, we fit each model (CART, linear regression, GAM, and MART) to one year's data (1997, for example), and used that model to predict cost for the other years (1996, 1998, and 1999, in this case).

Explore Predictive Ability of Other Functional Measures

The search for an ideal index set to predict cost occurred in two stages. First, we examined individual components. The main question was whether the components entered the model in the expected direction. More specifically, we fit a linear regression model predicting cost from the components of the motor and cognitive scale. We checked to see which, if any, of the components had positive coefficients--implying that greater functional independence increased cost. Such irregularities would flag further investigation of the data collection process for that component of the scale. We then might reconsider how or if it would be used in the index set. We also fit GAM to the components to look for nonlinear effects.

Second, we experimented with the subscales described in Stineman, Jette, et al. (1997b). These split out the standard motor index into dimensions reflective of different body areas and types of function.

Examine Stability Across Years

Our previous results were based on 1996 and 1997 data and did not give us much latitude for examining stability over time. We did verify, however, that the FRGs from the 1994 data predicted costs well in 1996 and 1997. With the addition of 1998 and 1999 data, we had the option of fitting models within each year and seeing how well they would do on three other years. We also had the ability to pool multiple years worth of data for RICs that were small and hence would not otherwise have had much out-of-sample predictive power.

Computational Design

Types of Models

Models that consistently predict cost well, in terms of the average square difference between the actual and predicted cost, across the various years and RICs, are termed "gold standard" models. Below, we list the types of models we fit, our reasons for fitting them, and how well they perform as gold standard models. Included with each of the methods is a two-dimensional visualization of the surface that each model fits to data. The data come from RIC 01 (Stroke) combining 1998 and 1999 data. The darkest regions of the plots show where the

model predicts the lowest cost for the motor and cognitive score combination. Since such visualization is limited to two dimensions, the plot intentionally excludes age.

OLS--Ordinary Least Squares. Linear models are fit with ordinary least squares regression. In a linear model, a fixed amount of change in an independent variable, anywhere along its scale, results in the same change in the prediction of the dependent variable. For example, a change in the motor score from 20 to 21 would decrease predicted cost by the same percentage as a change from 60 to 61. The coefficients of ordinary least squares report the increase in log-cost due to a unit increase in an individual component of the FIM.

The parallel contours in Figure 3.1 demonstrate this linearity. The darkest areas of the plot identify the regions with the least cost, and lighter areas indicate higher costs. The figure also shows that the strongest effect is due to the motor score. For a fixed value of the motor score, there is little change in cost for any value of the cognitive score. On the other hand, for patients with a particular cognitive score, their rehabilitation costs greatly depend on their motor score.

OLS affords a compact representation allowing for easy interpretation and diagnosis of the model. In particular, we looked for any coefficients that indicated that increases in functional ability tended to increase cost. We flagged these components for further investigation. Besides the interpretation, OLS is computationally inexpensive and often provides an accurate approximation to the relationship between log-cost and the functional measures. It would be an appropriate gold standard model if the assumption of a linear relationship between the independent variables and the dependent variable is true or approximately true.
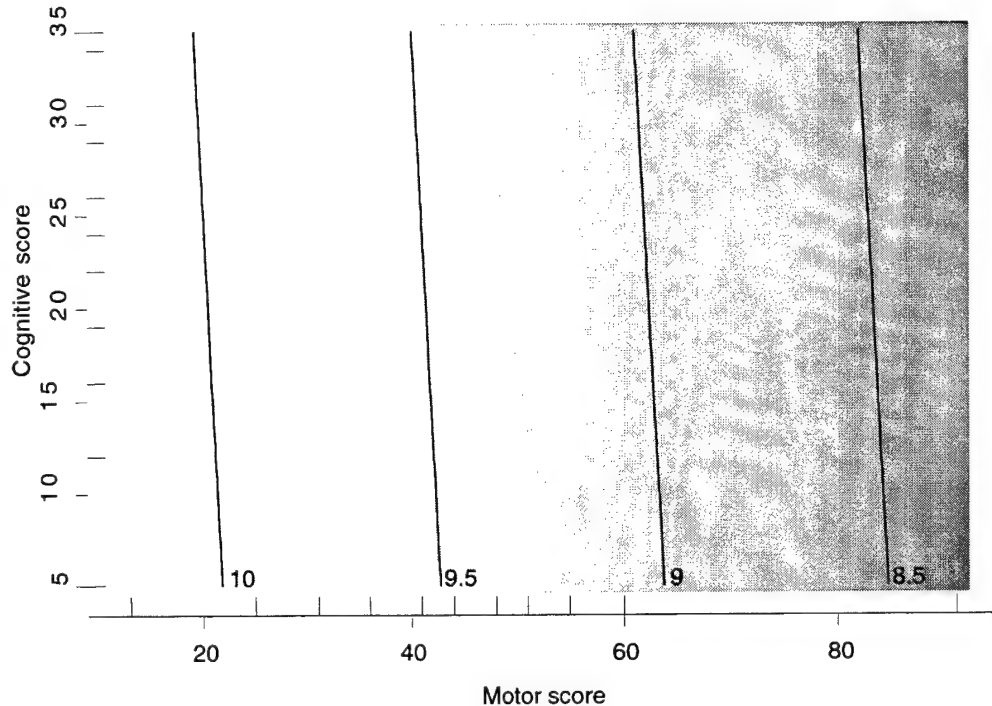
Figure 3.1--Linear Model

Generalized Additive Models (GAM). GAM permits slightly more flexible
relationships between the dependent and independent variables. GAM approximates
the relationship as a sum of smooth (rather than linear) functions of the
independent variables. This means that a change in motor score from 20 to 21
might decrease predicted cost by a different percentage than a change from 60 to
61. GAM does not model interactions; it only produces estimates of additive
effects. Because the relationship is assumed to be additive, the decrease in
predicted cost due to a change in motor score from 20 to 21 will be the same
regardless of the values of the other independent variables. The top two panels
of Figure 3.2 compare the linear model to GAM for predicting cost from motor
score. Although the two fits seem to agree closely, the GAM fit shows evidence
that the effect of motor score tapers off as motor score gets smaller. The MART
and CART plots are discussed in a later section.

Linear model

GAM

MART

CART
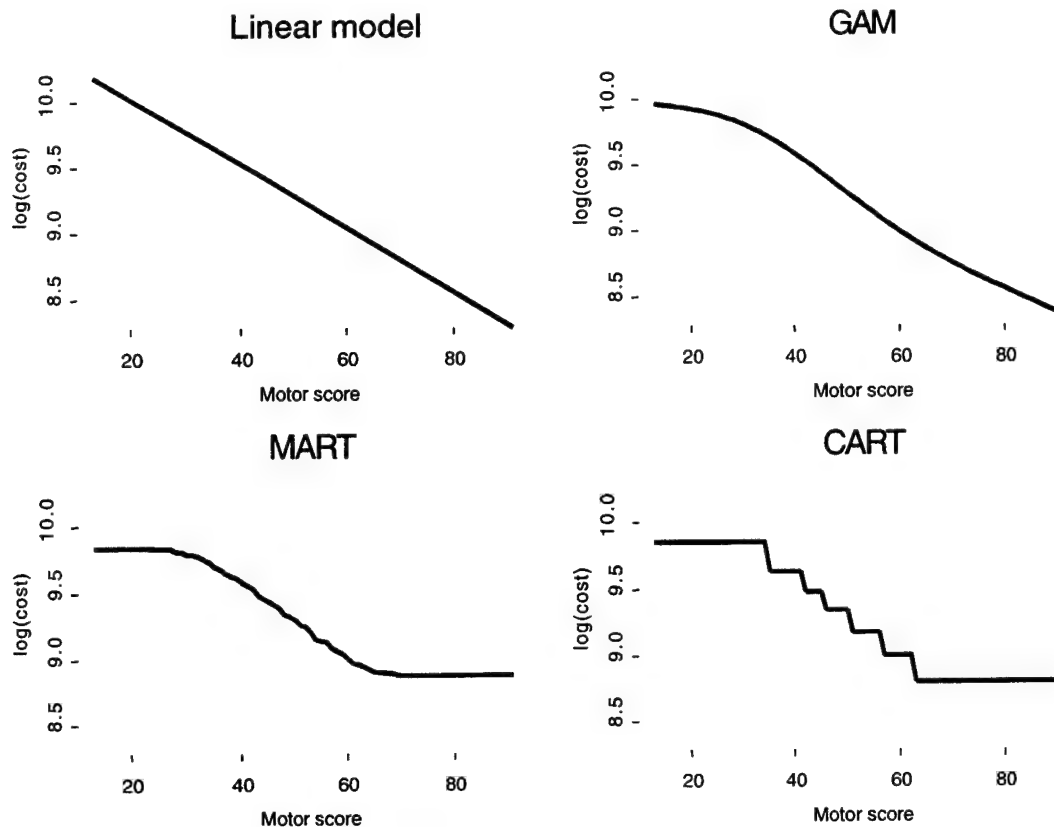
Figure 3.2--Comparison of Models for the Univariate Case

Although the additivity restriction may prevent the discovery of interaction effects in multivariate data, the benefits of additivity include easy computation and interpretation. To interpret GAM, we can plot, for each index, the value of the index versus the contribution it makes toward the log-cost estimate. We can then look for irregularities, saturation effects, and threshold effects. For example, we may learn that patients whose motor scores exceed a particular value have roughly constant cost, an example of a saturation effect. GAM uses more degrees of freedom than OLS but conserves them by imposing the additive constraint and restricting the additive components to be very smooth, spending roughly four degrees of freedom per predictor. GAM will also work well in small RICs.

Figure 3.3 shows the shape of the GAM fit. Clearly, GAM picks up curvature that the linear model cannot. It is still apparent that the motor score is the most influential. However, GAM also seems to pick up the fact that at extreme values on the cognitive scale the cost is slightly lower than for cognitive scores in the middle of the range.
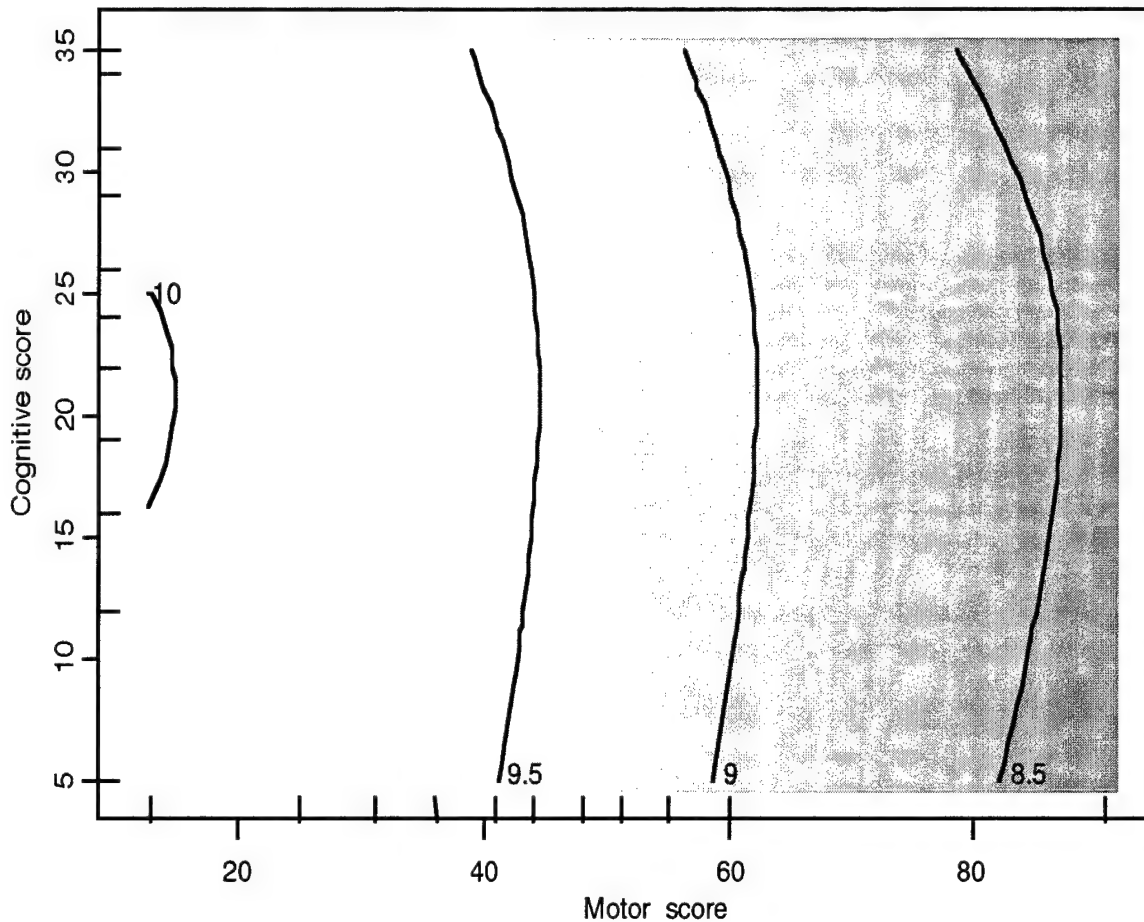
Figure 3.3--Generalized Additive Model

The cost of the additional flexibility is greater model complexity and variability. However, the same flexibility that makes GAM more complex can also make its predictions more accurate than the linear model's when the relationship between the dependent and independent variables is nonlinear.

Multiple Adaptive Regression Trees (MART). MART is a state-of-the-art statistical method. It is the most flexible and most complex of the models under consideration as a gold standard. Like GAM, it is nonparametric with the ability to find nonlinear relationships. However, it is also able to find interaction effects in the predictor variables.

The MART prediction is the sum of predictions from many simple CART models. The algorithm constructs the CART models sequentially in such a way that each additional CART model reduces prediction error. Since each CART model fits an interaction effect, the sum of many of them (hundreds to thousands) results in a

prediction model that permits complex, nonlinear relationships between the dependent and independent variables. We can control the depth of interaction effects that MART tries to capture by controlling the depth of the individual CART models. Unlike GAM and the linear model, MART is invariant to one-to-one transformations of the independent variables, so there is no need to search for the best scaling or logarithmic transforms of the predictors.

If cost varies in a non-additive way across motor score and cognitive scores, then MART might be able to capture this information and provide predictions that are more accurate than that of GAM. Figure 3.4 shows the shape of the MART fit. Like GAM, MART determines that the high cognitive scores have lower costs than the lower cognitive scores at a fixed motor score. Furthermore, MART shows that costs decrease much faster at the high cognitive scores for very low motor scores. This is a feature that the functional form of GAM cannot detect. When such effects are strong, MART would likely outperform GAM. This makes it a good candidate for the gold standard.

As with GAM, the additional complexity complicates interpretation. MART is difficult to interpret, and quantifying the number of degrees of freedom that it spends is difficult. However, some measures of variable influence and
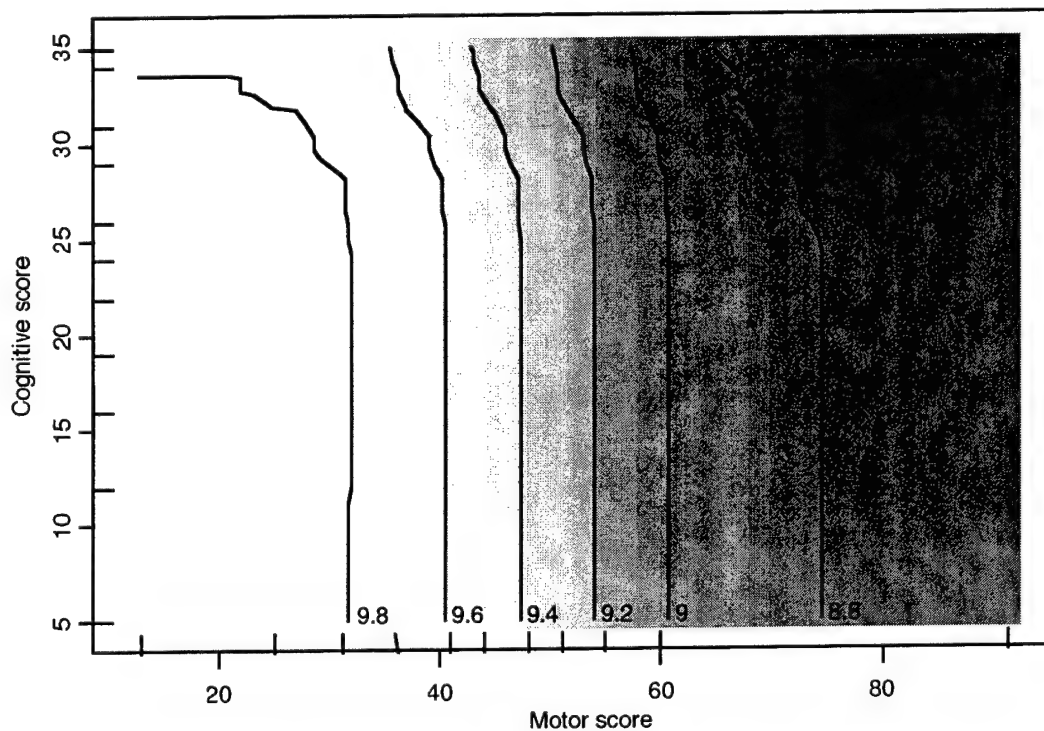
Figure 3.4--Multiple Adaptive Regression Trees

visualization tools are available for evaluating the predictor's rationale. It is not clear whether MART will always work well for very small RICs, but results show that it has been competitive with GAM.

Classification and Regression Trees (CART). CART is a well-known technique for building classification models. CART requires a dependent variable (here, log-cost), and it seeks to develop predictors of the dependent variable through a series of binary splits from a candidate set of independent variables (here, age, FIM motor score, and FIM cognitive score). As mentioned previously, CART partitions the data into two groups using the independent variables. Such a partition might separate patients with motor score exceeding 50 from those with motor score less than 50. CART chooses the variable on which to split the data and the value of the variable at which to split so that the new partitions are more homogeneous in terms of log-cost. The partition minimizes the square prediction error. CART then recursively splits each partition until it satisfies some stopping criterion.

Figure 3.5 shows how CART partitions in the data example. Figure 3.6 shows the dendrogram (tree) version of the plot. The findings are not unlike those of the previous analysis. We can still see that motor score is the primary effect, although at high motor scores, cognitive ability can be influential. Neither the linear model nor GAM can detect such interaction effects. However, the boundaries are abrupt and discontinuous. This is particularly noticeable in Figure 3.2. There, CART seems to be trying hard to fit a fairly smooth function, yet it is necessarily jagged. This kind of bias is likely to detract from the fit relative to GAM or MART.

By recursively partitioning the data, CART essentially fits interaction terms and thus can miss some main effects. CART has the pleasing theoretical property that as the sample size grows the prediction rule converges to the one that minimizes the expected prediction error. However, even in large samples CART can fail to fit curvature well (underfit) or can infer curvature where none exists (overfit). CART is a high-variance regression method, meaning that small fluctuations in the data set can produce very different tree structures and prediction rules. An early split will influence the shape of the tree and produce results that may be nonsensical. In practical use with large data sets, CART can produce a tree with many partitions, causing difficulty in interpretation and evaluation of the inferred rules.
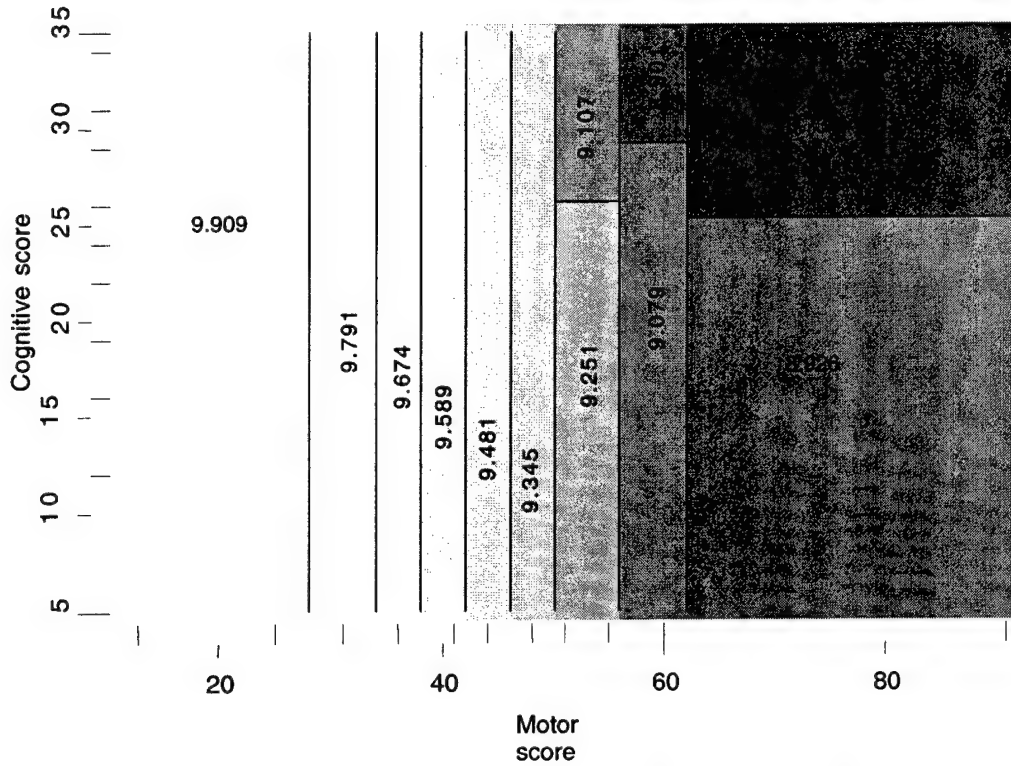
Figure 3.5--Classification and Regression Tree
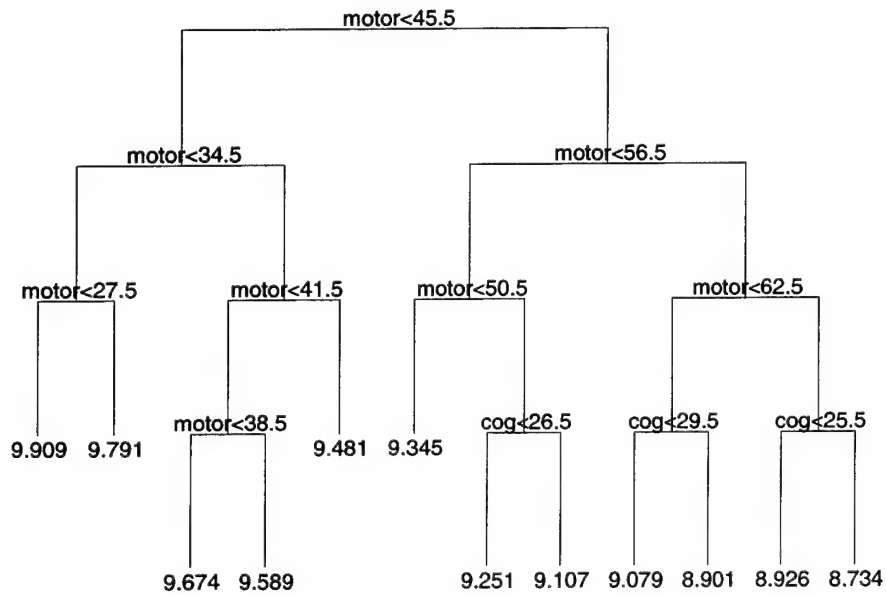
Figure 3.6--Dendrogram of the CART Model

Given these limitations, CART can still be a useful and powerful tool. The CART model offers the advantage of producing groups determined by ranges of the independent variables. Thus, it becomes easy to classify a new patient by comparing the values of the patient's set of independent variables with the ranges that define each of the CART determined groups. Our use of CART focused on three stopping criteria, all of which attempt to estimate the optimal number of partitions to generate from the data.

1. XVAL-CART: Standard cross-validation to minimize mean square error (MSE). CART's performance can be very sensitive to the number of partitions it produces. If there are too few partitions, patient groups with very different associated costs will not be separated. If there are too many partitions, the cost estimates will be unreliable and there will be difficulty in practical implementation of the payment formula.

We used tenfold cross-validation, the most widely used method, to estimate the optimal number of partitions. This method splits the data into ten groups containing equal numbers of patients. For each of the splits we construct a CART model on the other 90 percent of the observations and evaluate the performance of various tree sizes on that validation split. We then average the performance over the ten validation runs by tree size. We select the tree size with the lowest cross-validated mean square error to be the best tree size. We fit a final single CART model to the entire data set, stopping when the tree size reaches the tenfold cross-validation choice.

2. 1SD-CART: Stop when within one standard deviation of minimum MSE. Since we are working with fairly large data sets, it turns out that the tenfold cross-validation method can produce models with far too many splits. We needed to introduce "practical" considerations into the stopping criteria. Breiman et al. (1984, 78-80) recommend a more aggressive stopping rule to fix the number of partitions that corrects this situation. They suggest placing confidence bands around the cross-validated estimate of prediction error by tree size. Then choose the first node where prediction error is within one standard error of the minimum. This reduces the number of partitions, reduces the probability of overfitting, and could cause some more heterogeneous groups (in terms of log-cost) to be combined.

3. INT-CART: Interim report numbers of nodes. This algorithm simply fixes the number of nodes as that used in the project's interim report. This stopping

them. Indeed, when we initially fit CART models on all indices, the model fits were poor. We thought, however, that additive models might give useful information on the relative contributions of the individual items, and we attempted to apply both OLS and GAM using these items as predictors.

M13C5. Standard two scales (motor + cognitive). These were the standard indices that we had used for the interim report. This index set has only two terms--the sum of the 13 responses to the motor FIM items and the sum of the five responses to the cognitive FIM items.

M12C5. Standard, but excluding transfer to tub/shower in the motor score. This index set arose during the course of our investigation of individual items. As discussed further in the subsection "Results: Item level analysis," we found that patients with greater functional independence in transfer to tub/shower tended to cost more. Thus it is reasonable to believe that eliminating this item from the motor score may improve the prediction of cost.

This index set has only two terms--the sum of 12 responses to the motor FIM (transfer to tub excluded) and the sum of the five responses to the cognitive FIM.

M12C4. Standard, excluding transfer to tub in the motor score and comprehension in the cognitive score. This index set also arose during the course of our investigation of individual scores (see the Results subsection). In the cognitive FIM, increased functional independence on the comprehension component tended to increase cost. Consequently, we explore eliminating the comprehension item from the cognitive scale, although eliminating this item may be undesirable for reasons discussed below.

This index set contains only two terms--the sum of 12 of the responses to the motor FIM (transfer to tub excluded) and the sum of four of the responses to the cognitive FIM (comprehension excluded).

StJe3. Stineman and Jette, Activities of daily living (ADLs), mobility, standard cognitive. This and the following index set were proposed by Stineman et al. (1997a) as subscales of the cognitive and motor score that might relate to specific impairments. This index was found to describe dimensions of function within the large stroke RIC. Our question was whether additional information in these two indices could help to predict cost and to improve on the classification system, either in some or in all RICs. We set out to determine their potential contribution to cost prediction. Initially, we developed these

indices as defined by Stineman. However, once we determined that we preferred
the motor index without tub transfer (M12C5, above), we defined its
corresponding component mobility/transfer to exclude tub transfer as well.

This index set has three components--the sum of four of the mobility
components of the motor FIM (transfer to tub excluded), the sum of the eight
daily living components of the motor FIM, and the sum of all five of the
components of the cognitive FIM.

StJe5. Stineman and Jette, four motor scores (self-care, sphincter control,
transfer, locomotion) plus standard cognitive. This index set is a further
decomposition of the previous set. It subdivides the ADLs index into self-care
and sphincter and decomposes the mobility index into transfer and locomotion
sub-indices. Stineman and Jette found them to be dimensions of function in RICs
6 through 14, 17, 19, and 20.

This index set has five components--the sum of two of the transfer
components of the motor FIM (transfer to tub excluded), the sum of the two
locomotion components of the motor FIM, the sum of the two sphincter control
components of the motor FIM, the sum of the self-care components of the motor
FIM, and the sum of all five of the components of the cognitive FIM.

Fitting and Evaluation Periods

To validate the various estimators of the relationship between the indices
and log-cost we evaluated each method in terms of out-of-sample predictive
performance. The most important fits, of course, are the ones based on the most
recent data because they will determine the payment system. We can get an idea
of how well they perform by seeing how earlier years' fits perform on following
years' data. We initially tried fitting separate models for each year and seeing
how well they performed on all other years. This would yield 12 out-of-sample
fits/evaluations. We later improved on that by observing that some RICs (e.g.,
04, 11, 18, 19, 21) were quite small, and it might be advantageous to pool their
data. This led to experimenting with fitting periods 1996-1997 and 1998-1999.
Thus, the full set of fits and predictions is described by the appearance of "x"
in Table 3.7.

- 55 -

Table 3.7

Combination of Fitting and Evaluation Periods Examined

	Evaluation Period			
Fitting Period	**1996**	**1997**	**1998**	**1999**
1996		x	x	x
1997	x		x	x
1998	x	x		x
1999	x	x	x	
1996-1997			x	x
1998-1999	x	x		

Results

Item-Level Analysis

We ran OLS with individual subscales (eating, walking, etc.). We wanted to know whether the individual items appeared to influence costs in the expected direction: Higher FIM scores should mean lower costs. OLS with log-cost would be the easiest method to interpret. If the estimated coefficients were positive and we wound up paying less for cases with less function, it would provide incentives that many clinicians would find unacceptable. Although there will be some cases where patients with lower functional independence can profit less from extended therapy than can patients with more independence, there will be other lower-functional-independence cases where the opposite is true. To pay less for the whole group might provide access problems for the less independent patients who can benefit from the therapy.

Randomness alone would produce numerous positive regression coefficients: There are 18 individual components and 21 RICs, and some coefficients cannot be measured precisely; so we expect a number of small t-statistics that could be on either side of zero. However, we have good power to detect if a given effect is consistently positive. If the pattern of positive signs persists for all four years of data and for several RICs within each year, we would have some confidence that we had found an anomalous item.

Table 3.8 shows, for the OLS regressions, how many RICs had a positive sign within each of the data sets across the 21 RICs, and how many of these coefficients had t-statistics greater than 1.0. The unmistakable patterns are that both tub transfers and comprehension often have the wrong sign in OLS regressions, implying that costs were higher when the functional independence measure was higher.

Table 3.8

Component Regressions:
Occurrences of Positive Regression Coefficients in 21 RICs

Variable	Positive OLS Coefficients						OLS t-statistic ≥ 1.0					
	1996	1997	1998	1999	96-97	98-99	1996	1997	1998	1999	96-97	98-99
Comprehension	19	15	19	19	20	20	13	11	15	15	15	16
Expression	7	4	6	7	4	4	3	1	2	1	1	3
Social interaction	14	10	7	9	11	7	6	3	1	3	5	3
Problem solving	3	4	4	7	3	4	0	1	2	2	0	2
Memory	5	5	4	2	4	1	3	2	0	1	1	0
Eating	1	0	1	0	0	1	1	0	0	0	0	0
Grooming	9	11	13	12	8	12	5	7	9	7	6	10
Bathing	5	2	2	1	2	2	0	1	1	0	0	0
Dress upper body	13	10	12	14	11	15	3	3	7	10	6	10
Dress lower body	0	3	1	1	1	0	0	1	0	0	1	0
Toileting	0	0	1	1	0	1	0	0	0	0	0	0
Bladder	1	1	1	3	1	0	0	0	0	0	0	0
Bowel	11	6	14	12	7	13	4	5	3	5	6	6
Transfer to bed	2	2	0	0	1	0	0	1	0	0	1	0
Transfer to toilet	2	1	0	2	0	0	1	0	0	0	0	0
Walking	1	0	0	0	0	0	0	0	0	0	0	0
Stairs	5	3	5	4	2	4	2	2	2	2	2	1
Transfer to tub	17	19	21	18	20	20	14	13	18	16	13	18

We believe that the perverse effect of transfer to tub is due to the fact that the response depends on the situation being scored--either tub or shower and with or without assistive devices. The UDSmr question-and-answer manual says: "It may be that a subject's score goes down as he/she no longer requires the use of some assistive device" (p. 31). It is likely that patients who do not use a device at admission score worse than patients who do, so the FIM item provides only a situational measure of independence rather than an absolute measure. Further, transfer to tub/shower is frequently not assessed at admission and instead is rated "least independent" regardless of what the patient's ability actually is.

We have no similar rationale for the comprehension results. It may be that this finding reflects only the fact that many hospitals do more for patients who understand what is happening. Without making a judgment about whether this is appropriate behavior on average, one can still see that eliminating this item raises issues related to incentives and fairness. Removing this item from the index may remove the incentive to treat patients with low comprehension or

penalize hospitals that make additional efforts to treat them. As shown below, the cognitive scale is used only occasionally in the definition of CMGs. When it is used, the other four items in the cognitive scale determine the direction of the cognitive effect, so that a higher cognitive score results in a lower payment. To confirm that OLS was not overlooking important nonlinear effects, we also looked at plots of the marginal contributions of each component, as estimated by GAM. Although log-cost did not always smoothly decline as each component advanced along its seven-point scale, there were only two items where the relationship was perverse. These were the same ones that showed up in the linear models: comprehension and transfer to tub.

One other piece of information that we wanted to examine was the mode used in the walk/wheelchair item. We experimented with a wheelchair dummy variable and an interaction between the wheelchair dummy and the FIM walk/wheelchair response in our OLS and GAM fits. We found that, in most RICs, wheelchair patients cost more than expected given their functional scores, and that locomotion score is less important for those in wheelchairs than for those who walk. The net effect of wheelchair alone was quite small. However, in the two RICs that have the most wheelchair people (RICs 4 and 10), neither wheelchair functional status nor the wheelchair indicator is significant. Thus, adding these variables will not result in a substantial improvement in the prediction of cost.

Selecting a Gold Standard Model

The gold standard model does two things. First, it helps us understand how well CART is doing by giving us a measure of attainable residual standard deviation with which we can compare the residual standard deviation we get from CART. Second, in a simulation exercise it enables us to assess the prediction bias for various combinations of demographic and hospital factors.

MART and GAM are the candidates for gold standard status. We have theoretical reasons to prefer MART. It is extremely flexible, and it detects interactions. However, its prediction formula is rather unwieldy. Also, some RIC sample sizes are small, and it may be that without forcing some structure, one effectively fits too many parameters and gets a model that does not extrapolate very well. On the other hand, GAM uses fewer degrees of freedom and produces a curve to describe the effects of each input variable, so it is a little easier

to decide whether the GAM fits make clinical sense. Without a clear a priori winner, we decided to perform our computations on both GAM and MART.

Knowing that CART would not produce reasonable models with component scores alone, we chose not to work further with the components at this point. We fit all combinations of models and remaining indices (six types of models, five types of indices, six fitting periods). We looked at out-of-sample root mean square prediction error (RMSE) as a measure of quality of fits. Aggregate RMSEs across RICs are provided in Table 3.9.

Table 3.9

Root Mean Square Errors Among Candidate Gold Standard Models

Fit Year	Eval. Year	Const.	M13C5 GAM	M13C5 MART	M12C5 GAM	M12C5 MART	M12C4 GAM	M12C4 MART	StJe3 GAM	StJe3 MART	StJe5 GAM	StJe5 MART
96	97	.541	.474	.474	.473	.473	.473	.473	.470	.470	.467	.466
	98	.545	.480	.480	.479	.479	.479	.479	.475	.475	.473	.472
	99	.546	.483	.484	.482	.482	.482	.482	.478	.479	.476	.476
97	96	.535	.468	.468	.467	.467	.467	.467	.464	.464	.462	.461
	98	.545	.480	.479	.479	.478	.479	.478	.475	.474	.472	.471
	99	.546	.483	.483	.482	.482	.482	.481	.478	.478	.476	.475
98	96	.535	.469	.469	.468	.468	.468	.467	.465	.465	.462	.462
	97	.541	.475	.474	.474	.473	.473	.473	.471	.470	.467	.466
	99	.546	.482	.482	.481	.481	.481	.481	.477	.476	.475	.474
99	96	.535	.469	.469	.468	.468	.468	.468	.465	.465	.463	.463
	97	.541	.475	.474	.474	.473	.474	.473	.471	.471	.468	.467
	98	.545	.480	.479	.479	.478	.478	.478	.475	.474	.472	.471
96-97	98	.545	.480	.479	.479	.478	.479	.478	.475	.474	.472	.471
	99	.546	.483	.482	.482	.481	.481	.481	.478	.477	.476	.475
98-99	96	.535	.469	.468	.468	.467	.468	.467	.465	.464	.462	.461
	97	.541	.475	.474	.473	.473	.473	.472	.470	.469	.467	.466

The RIC constant column fits means to RICs (i.e., one FRG per RIC). Thus it is the within-RIC standard deviation of the log of the cost of cases, using a case-weighted average across RICs. It measures the amount of standard deviation that might be explained by defining FRGs within each RIC. Then, for each index set, we show RMSEs for GAM and MART.

There are several interesting things to observe. First, MART sometimes does slightly better than GAM, but they do about equally well. Second, the index without transfer to tub (M12C5) does slightly better than the index with it (M13C5) in all combinations of fitting year and prediction year. Comparing M12C5 to the similar model without the comprehension item (M12C4), we find that dropping comprehension improves prediction in only six of the 32 predictions that we evaluated. Third, both GAM and MART seem to be able to make use of

- 59 -

additional index information. RMSE goes down as the number of indices goes up, and the RMSE is lowest for the most disaggregated set of indices--StJe5. Finally, the RMSEs are all large, even for StJe5. About 15 percent of the standard deviation, or 25 percent of the variance, is explainable. But we cannot do better than that by creating FRGs. Case-level costs are inherently unpredictable.[1]

For completeness, we include the corresponding table of R-squareds as Table 3.10. The table shows that using RIC alone to predict cost explains about 16 percent of the variance. Adding the FIM and age variables increases the percentage explained to about 38 percent. We think this is near the upper bound of what can be explained in modeling case-level cost data.

We also looked at reductions in standard deviation within each RIC. Percentage reductions varied from about 20 percent for stroke (RIC 01) to about 10 percent for the three orthopedic RICs (07, 08, 09). These orthopedic RICs are substantially more homogeneous in cost than other RICs, so that despite the fact that we predict a smaller fraction of the variance in these RICs, they have RMSEs that are among the lowest of all other RICs. Figures 3.7 and 3.8 show what was typical of most RICs: MART and GAM perform about equally well, but MART does a little better across prediction periods. Performance in the smaller RICs was similar (see Figures 3.9 and 3.10), although sometimes GAM did a little better. But in no case did GAM seem to dominate MART across prediction periods.

---

[1] The payment system, of course, also exploits the variance across RICs in cost. About 34 percent of the total variance in the wage-adjusted cost of cases discharged to the community is predicted by the FRG system and 37 percent by our gold standard models.

Table 3.10

R-Squareds Among Candidate Gold Standard Models

Fit Year	Eval Year	Const	M13C5		M12C5		M12C4		StJe3		StJe5	
			GAM	MART	GAM	MART	GAM	MART	GAM	MART	GAM	MART
96	97	.16	.35	.36	.36	.36	.36	.36	.37	.37	.38	.38
	98	.15	.34	.34	.35	.34	.35	.35	.36	.36	.36	.36
	99	.15	.33	.33	.33	.33	.34	.34	.34	.34	.35	.35
97	96	.17	.36	.37	.37	.37	.37	.37	.38	.38	.38	.39
	98	.15	.34	.34	.35	.35	.35	.35	.36	.36	.36	.37
	99	.15	.33	.33	.34	.34	.34	.34	.34	.35	.35	.35
98	96	.17	.36	.36	.37	.37	.37	.37	.38	.38	.38	.38
	97	.16	.35	.36	.36	.36	.36	.36	.37	.37	.38	.38
	99	.15	.33	.33	.34	.34	.34	.34	.35	.35	.35	.36
99	96	.17	.36	.36	.37	.37	.37	.37	.37	.37	.38	.38
	97	.16	.35	.35	.36	.36	.36	.36	.36	.37	.37	.38
	98	.15	.34	.35	.35	.35	.35	.35	.36	.36	.36	.37
96-97	98	.15	.34	.34	.35	.35	.35	.35	.36	.36	.36	.37
	99	.15	.33	.33	.34	.34	.34	.34	.34	.35	.35	.35
98-99	96	.17	.36	.36	.37	.37	.37	.37	.38	.38	.38	.38
	97	.16	.35	.36	.36	.36	.36	.36	.37	.37	.38	.38
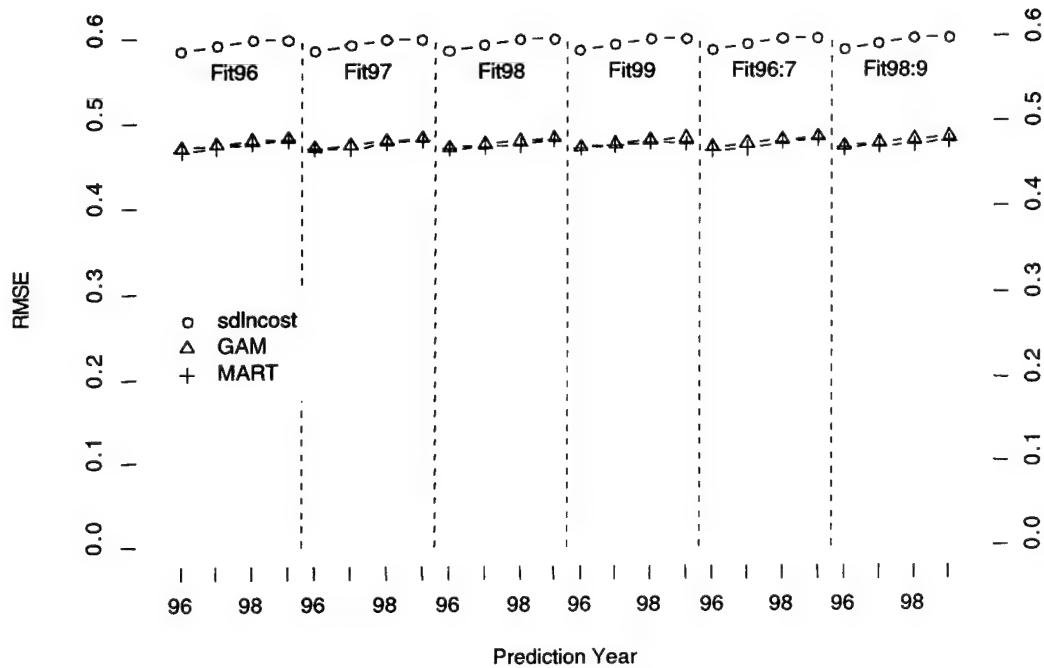
Figure 3.7--RMSEs, by Fit and Prediction Years: RIC=01, Index=StJe5
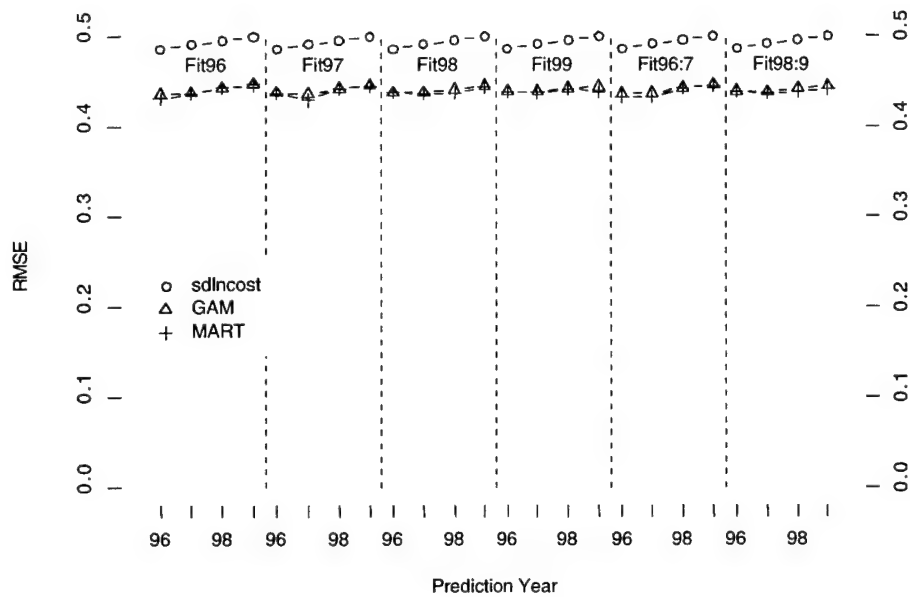
Figure 3.8--RMSEs, by Fit and Prediction Years: RIC=07, Index=StJe5
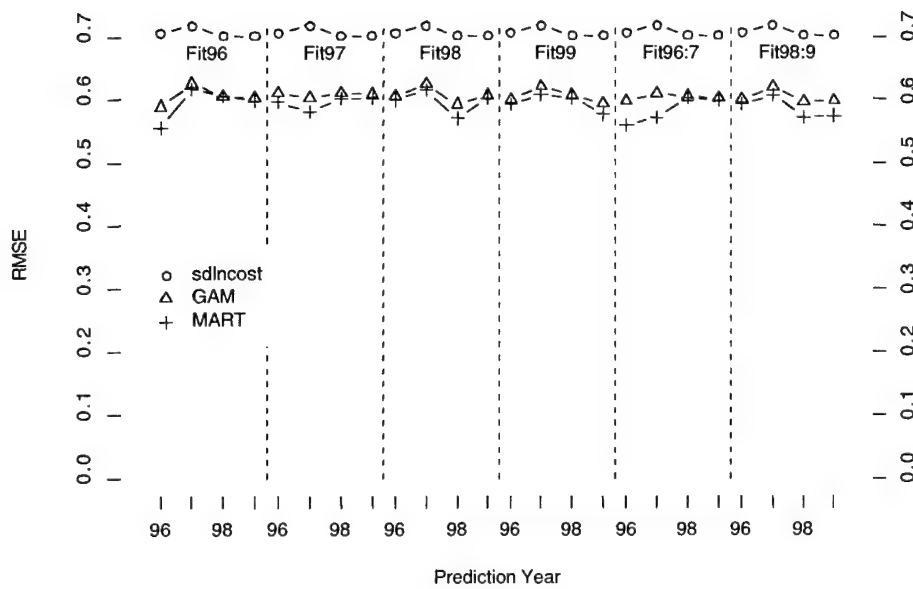
Figure 3.9--RMSEs, by Fit and Prediction Years: RIC=04, Index=StJe5
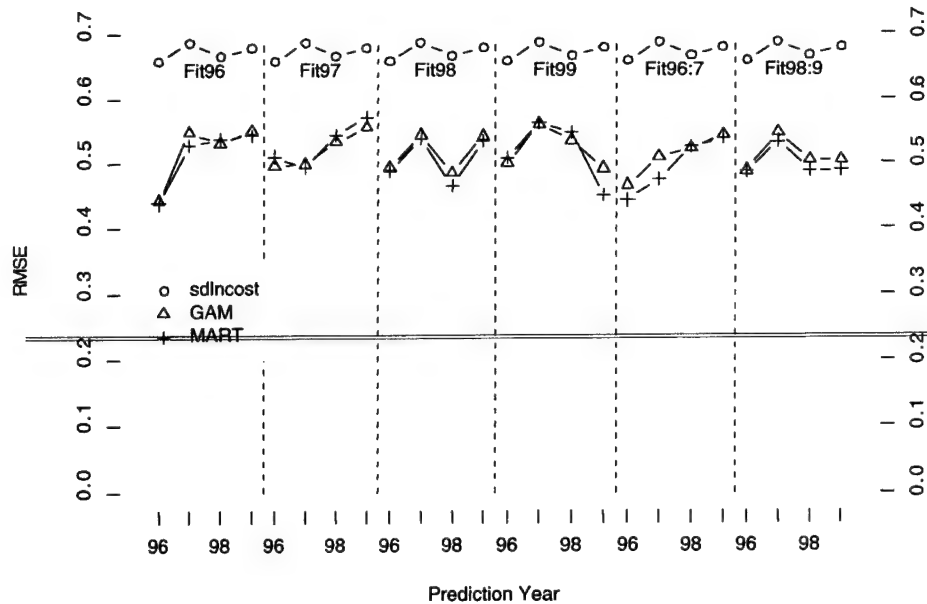
Figure 3.10--RMSEs, by Fit and Prediction Years: RIC=18, Index=StJe5

In summary, we saw that MART seemed to be a little better than GAM for
many RICs, validating the observations we made above at the aggregate level.
Also, it appears that MART is able to make use of the more complicated index
information because the RMSEs decrease when we use the finer subscale
components. For the purpose of determining the percentage of variance explained,
MART with index set StJe5 is an appropriate gold standard, and we use this in
our evaluations below. However, we also recognize that it may be too strong a
gold standard--there are times when we want to stay with a given index set and
ask how much better we might do by varying only the functional form. In such
cases, based on the observation from Table 3.9 that MART performs at least as
well as GAM for every choice of index set, we use MART with the specified index
set.

Evaluation of CART Models

We have shown above that we achieve the best prediction using MART and the
StJe5 index set. But this does not lead to a simple payment system, and it may
not meet the definition of patient groups found in the law. We would need a
complex computer program to evaluate the formula. This is not compatible with
the design criteria of the payment system. We therefore employ CART to produce
simple, understandable patient groups.

We introduce the alternative CART models we considered by reviewing the
results for the index set that we believe is best, M12C5. Table 3.11 shows RMSEs

Table 3.11

Performance of Alternative CART Models: Index = M12C5

Fit Year	Eval Year	Within-RIC Standard Deviations						Percentage of SD Explained by StJe5[a]			Percentage of SD Explained by M12C5[b]			Number of Nodes		
		Const	INT	1SD	XVAL	StJe5-MART	M12C5-MART	INT	1SD	XVAL	INT	1SD	XVAL	INT	1SD	XVAL
96	97	.541	0.479	0.480	0.477	0.466	0.474	82.7	81.3	85.3	92.5	91.0	95.5	104	95	218
	98	.545	0.485	0.486	0.483	0.472	0.480	82.2	80.8	84.9	92.3	90.8	95.4			
	99	.546	0.489	0.489	0.486	0.476	0.483	81.4	81.4	85.7	90.5	90.5	95.2			
97	96	.535	0.473	0.473	0.471	0.461	0.468	83.8	83.8	86.5	92.5	92.5	95.5	104	97	244
	98	.545	0.485	0.485	0.482	0.471	0.480	81.1	81.1	85.1	92.3	92.3	96.9			
	99	.546	0.488	0.488	0.486	0.475	0.483	81.7	81.7	84.5	92.1	92.1	95.2			
98	96	.535	0.474	0.473	0.471	0.462	0.469	83.6	84.9	87.7	92.4	93.9	97.0	104	123	333
	97	.541	0.479	0.479	0.477	0.466	0.475	82.7	82.7	85.3	93.9	93.9	97.0			
	99	.546	0.487	0.486	0.484	0.474	0.482	81.9	83.3	86.1	92.2	93.8	96.9			
99	96	.535	0.474	0.474	0.472	0.463	0.469	84.7	84.7	87.5	92.4	92.4	95.5	108	126	325
	97	.541	0.480	0.479	0.477	0.467	0.475	82.4	83.8	86.5	92.4	93.9	97.0			
	98	.545	0.484	0.483	0.481	0.471	0.480	82.4	83.8	86.5	93.8	95.4	98.5			
96-97	98	.545	0.485	0.483	0.482	0.471	0.480	81.1	83.8	85.1	92.3	95.4	96.9	104	142	398
	99	.546	0.488	0.486	0.485	0.475	0.483	81.7	84.5	85.9	92.1	95.2	96.8			
98-99	96	.535	0.474	0.471	0.470	0.461	0.469	82.4	86.5	87.8	92.4	97.0	98.5	108	180	483
	97	.541	0.479	0.477	0.475	0.466	0.475	82.7	85.3	88.0	93.9	97.0	100.0			

[a] The percentage of standard deviation explained by the model, where 0 equals the constant term model, and 100.0 equals the gold standard (MART with index set StJe5).

[b] The percentage of standard deviation explained by the model, where 0 equals the constant term model, and 100.0 equals MART with index set M12C5.

for the M12C5 model. The further to the right, the more FRGs in the CART model. The RIC constant column fits means to RICs (i.e., one FRG per RIC) and repeats the data from Table 3.9. The interim report used the 1SD rule with some adjustments; it performs similarly to the 1SD rule applied here. The XVAL column shows how well CART does with its standard cross-validation stopping rule, which tends to produce more than twice the terminal nodes as 1SD.

The main observation is that the CART models traverse a substantial fraction of the distance between the constant model and the gold standard. The CART FRGs explain more than 80 percent of the explainable standard deviation. If we restrict comparison to MART with the same index set (M12C5), we see that the CART FRGs explain more than 90 percent of the standard deviation. The 1SD model, which has less than half the nodes of XVAL, explains almost as much as XVAL.

Table 3.12 compares the performance of the alternative CART models relative to the gold standard for all of the indices we considered. It shows once again that M12C5 outperforms M13C5. Considering the results for the INT models, which force an equal number of nodes, we notice that M12C5 does better per node than either StJe3 or StJe5. M12C4 performs slightly better than M12C5, but the improvement is much smaller than that between M13C5 and M12C5.

Having settled upon CART for the payment formula, we are left to decide between the candidate indices. The criteria are quality of fit and parsimony. In the interim report we used 1SD-CART and 92 nodes. If we use 1SD-CART with the multiple indices in StJe3 and StJe5, CART produces far too many nodes. Raw 1SD-CART numbers of nodes in 1999 have 186 nodes for StJe3 and 201 nodes for StJe5 with almost no improvement in RMSE. If we fit with the interim report number of nodes, we get RMSEs for StJe3 and StJe5 that are larger.

In CART, the index with transfer to tub (M13C5) does noticeably worse than the index without this item in many years and never does substantially better. This is similar to our findings with GAM and MART. Because we believe this item does not measure an absolute level of function, we recommended to HCFA that this item not be used in creating FRGs. So, the choice seems to be between M12C5 and M12C4.

Table 3.12

Performance of Alternative CART Models: Percentage of SD Explained

Fit Year	Eval Year	Index=M12C5			Index=M12C4			Index=M13C5			Index=StJe3			Index=StJe5		
		INT	1SD	XVAL	INT	1SD	XVAL	INT	1SD	XVAL	INT	1SD	XVAL	INT	1SD	XVAL
96	97	82.7	81.3	85.3	82.7	82.7	85.3	78.7	77.3	80.0	78.7	82.7	86.7	80.0	84.0	88.0
	98	82.2	80.8	84.9	82.2	80.8	84.9	78.1	74.0	78.1	79.5	82.2	87.7	80.8	83.6	87.7
	99	81.4	81.4	85.7	82.9	81.4	85.7	78.6	77.1	81.4	80.0	82.9	87.1	80.0	82.9	87.1
97	96	83.8	83.8	86.5	83.8	83.8	86.5	81.1	81.1	85.1	79.7	83.8	87.8	79.7	85.1	86.5
	98	81.1	81.1	85.1	81.1	82.4	85.1	79.7	79.7	83.8	79.7	82.4	86.5	77.0	83.8	83.8
	99	81.7	81.7	84.5	81.7	81.7	85.9	80.3	78.9	84.5	78.9	83.1	87.3	77.5	84.5	84.5
98	96	83.6	84.9	87.7	83.6	84.9	87.7	82.2	82.2	86.3	80.8	84.9	87.7	80.8	86.3	89.0
	97	82.7	82.7	85.3	82.7	84.0	86.7	81.3	81.3	85.3	78.7	84.0	86.7	80.0	85.3	88.0
	99	81.9	83.3	86.1	81.9	83.3	86.1	79.2	80.6	84.7	79.2	84.7	87.5	79.2	84.7	88.9
99	96	84.7	84.7	87.5	84.7	86.1	88.9	83.3	83.3	86.1	80.6	86.1	88.9	80.6	86.1	90.3
	97	82.4	83.8	86.5	83.8	85.1	87.8	81.1	82.4	85.1	79.7	85.1	87.8	81.1	86.5	90.5
	98	82.4	83.8	86.5	82.4	83.8	86.5	79.7	81.1	85.1	79.7	85.1	87.8	81.1	86.5	90.5
96-97	98	81.1	83.8	85.1	81.1	83.8	85.1	79.7	82.4	85.1	79.7	86.5	89.2	81.1	87.8	90.5
	99	81.7	84.5	85.9	81.7	84.5	85.9	80.3	83.1	85.9	80.3	87.3	90.1	80.3	87.3	90.1
98-99	96	82.4	86.5	87.8	82.4	86.5	87.8	82.4	83.8	86.5	79.7	86.5	89.2	79.7	87.8	90.5
	97	82.7	85.3	88.0	82.7	85.3	88.0	81.3	84.0	86.7	80.0	86.7	89.3	80.0	89.3	92.0

Note: SD = Standard deviation.

The argument for going to M12C4 is that comprehension seems to work opposite to the standard cognitive scale in which it is embedded. After a stopping rule is fixed, dropping comprehension from the index produces a slightly better prediction in some years. However, when the full cognitive scale is used, the other four items in the cognitive scale determine the direction of the cognitive effect, so that a higher cognitive score results in a lower payment. We could eliminate splits that contradict this general result if they were to occur. If we take the comprehension item out of the index, however, the system will provide no extra incentives to treat patients with lowered comprehension. If some hospitals did spend extra to treat such patients, they would not be compensated for such extra resources. Further, the reduced cognitive scale does not increase the frequency with which FRGs are defined by cognitive function. On the other hand, if this observed decrease in cost with declines in comprehension really represents the current pattern of best care, our system should reflect it. The improvements in predicting cost are so slight that it seemed to us that the decision should be based on clinical judgment about what should be paid for. Based on the advice of our TEP, we recommend

keeping the comprehension item in the cognitive score. We will further explore this in phase II.

Cost Patterns

We wanted to understand the marginal contribution of motor and cognitive scores to the estimated log-cost. OLS coefficients provide such marginal estimates, but they enforce linear effects. GAM provides marginal estimates and allows arbitrary curvature. We attempted to understand the patterns of fit by graphing our GAM-M12C5 fits versus the motor and cognitive scales. Because the GAM fits were almost as good as MART's, we thought this would give an accurate portrayal of the cost versus scale relationships. Those graphs are shown in Figures 3.11 through 3.17 for a representative selection of RICs. We computed and examined these graphs for fitting year 1999 and for pooled 1998 and 1999 data.
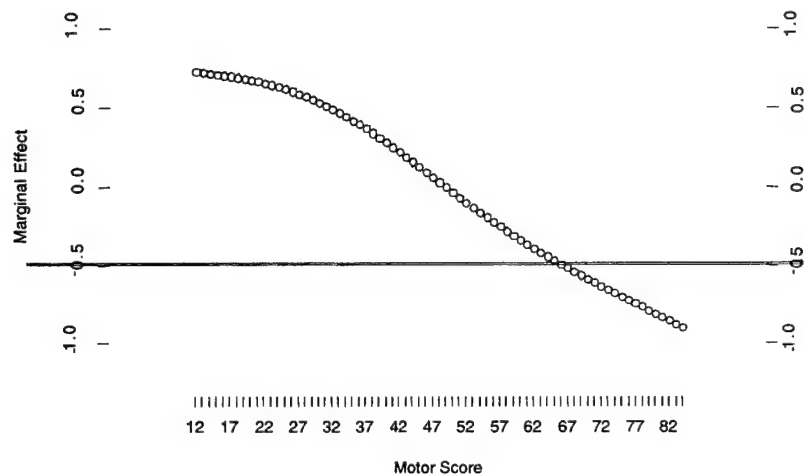
Figure 3.11--GAM Motor Scale Fits: RIC=01, Fityear=99
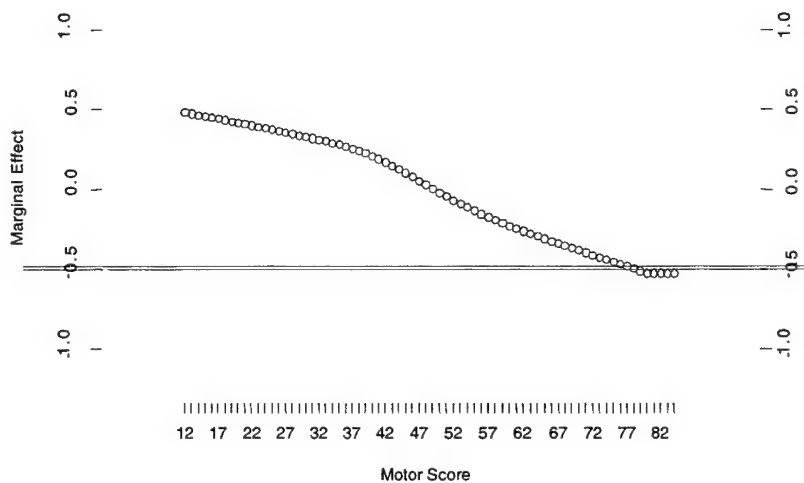
Figure 3.12--GAM Motor Scale Fits: RIC=08, Fityear=99
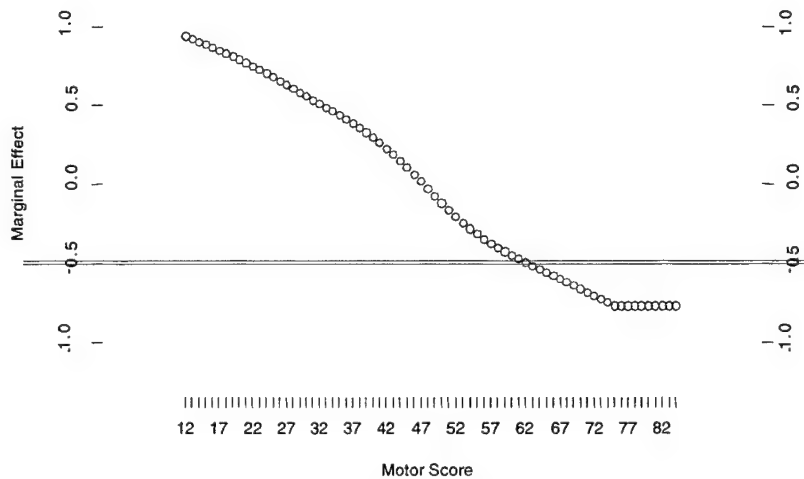
Figure 3.13--GAM Motor Scale Fits: RIC=19, Fityear=98,99
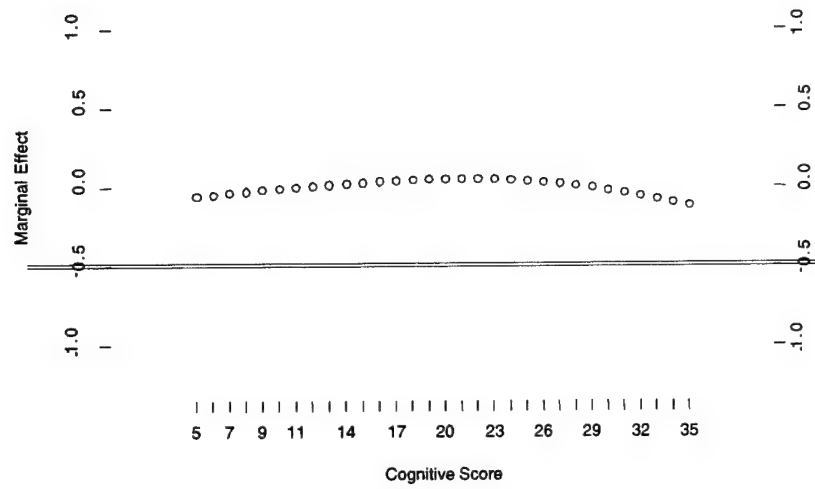
Figure 3.14--GAM Cognitive Scale Fits: RIC=01, Fityear=99
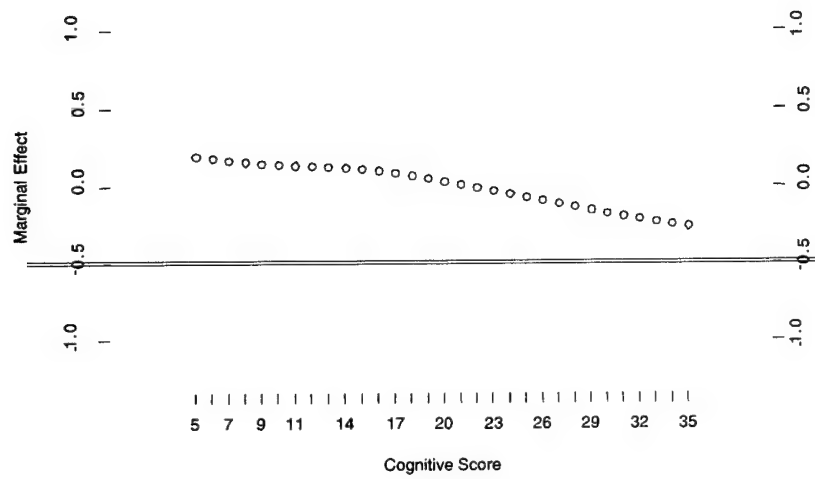
Figure 3.15--GAM Cognitive Scale Fits: RIC=02, Fityear=99
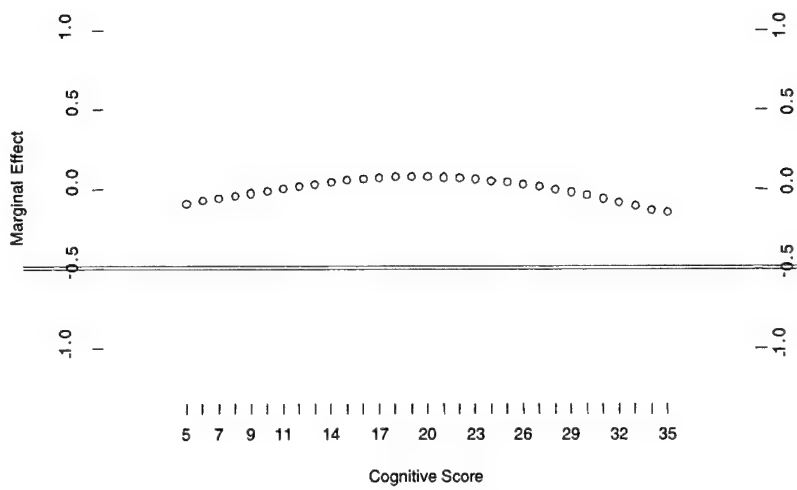
Figure 3.16--GAM Cognitive Scale Fits: RIC=08, Fityear=99
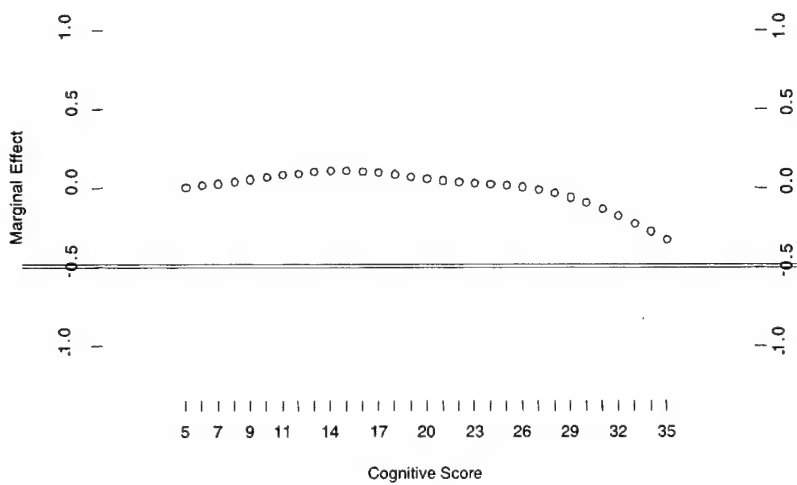
Figure 3.17--GAM Cognitive Scale Fits: RIC=18, Fityear=98,99

The plots are centered at zero and are uniformly scaled to span the range of effects for all RICs. One can see that the motor effects are always strong and sloping in the expected direction (larger scores yield lower costs). The cognitive effects tend to be much smaller, very close to zero. Also, higher scores are often associated with higher costs. Since CART attempts to replicate these patterns, it will largely split on motor scores, and hence the FRGs will simply reflect the motor score effect.

The GAM curves show cost as a function of both the motor and cognitive scales. If there were discontinuities in these curves, one would expect CART to discover them and to explain a lot of the variation. But the cost curves are continuous. At best, CART must approximate a smooth curve by a (small) series of discrete jumps. This might lead to the expectation of certain instabilities in CART's choice of cut-points, and that different data sets will indeed produce trees where cut-points differ. However, CART will find the steepness with respect to motor scores and should be expected to produce a lot of motor score splits.

Cost is strongly influenced by motor scores and in the expected direction. Except at certain motor score extremes, where there are few data, the higher the motor score, the lower the cost. On the other hand, the cognitive effects (Figures 3.14-3.17) are relatively flat and frequently not monotone. This is true for both indices M12C4 and M12C5. If we were to ask CART to discover the cost pattern, it might produce FRGs that are not monotone, which we think could pose problems for a payment system. The appropriate public policy decision might be to never lower payments for patients admitted with lower functionality (i.e., to develop monotone cost curve estimates).

To see how far some of the GAM models were from monotone fits, we tried fitting the closest monotone function to each of the GAM patterns in a least-squares sense and seeing how much of a difference this made in the percentage of standard deviation explained. The RMSE increase, to three significant digits, was zero. Thus, the monotone cost curves fit the data as well as the unconstrained cost curves.

Summary

Among the indices we examined, the standard cognitive index and a motor index that excludes transfer to tub do as well or better than the

alternatives we examined, although a cognitive index that excludes the comprehension score also performs well and deserves further consideration.

Our current recommendation is to use M12C5 with 1SD-CART, which achieves more than 80 percent of the maximum possible reduction in standard deviation. We evaluated RMSEs and found that M12C5 outperformed the standard motor and cognitive scales. It outperformed the expanded index set lists (StJe3, StJe5) in CART models where we constrained the number of nodes, but it was about equal to and slightly worse than M12C4 in its overall performance.

We find that the patterns of variation are described by a strong relationship between motor score and cost (higher motor scores are associated with lower cost) and a weak relationship between cognitive scores and cost. The fitted curves do not appear to be far from monotone approximations that enforce an inverse relationship between cost and FIM scores. This implies that that data will support a "monotone" payment scheme in which higher FIM scores never lead to higher payments--perhaps a politically desirable situation.

RECOMMENDED FRGS

Selection of FRGs

The preceding subsection supports our intention to develop FRGs through 1SD-CART models using the index M12C5. In obtaining FRGs, we wanted to accommodate the following considerations:

- If it looked as if we could improve the fits by the addition of a node or two, we did so.
- Fits should be monotone, decreasing in both the cognitive and motor indices. That is, if the FIM score were to increase by one, then the resulting predicted cost should not increase.
- The number of nodes should be manageable, roughly 100. For administrative simplicity, we did not wish to create a large number of groups. In addition, we did not want to create groups characterized by very small intervals of motor or cognitive scales for fear it would encourage upcoding.
- Groups that differ on a single factor (i.e., adjacent nodes of a tree) should differ "significantly" in payment amount.

In addition, we wanted to implement these in a formal algorithm that could be judged on its own merits devoid of subjective considerations. After considerable experimentation, we arrived at the following sets of rules:

- Pool the data for 1998-1999 where RIC sample sizes are less than 1,000; this results in pooling for RICs 04, 11, 18, 19, and 21. Use 1999 data only for the other RICs.
- Fit the 1SD-CART tree within each RIC.
- Consider adding nodes to the 1SD trees where R2 (equivalently, RMSE) will improve significantly. Look at traces of R2 versus number of nodes, and look to see where the addition of a node would improve R2 by 4 percent or more; add that node, and repeat this step until additional node contributions are found to be less than 4 percent.
- Produce tables describing trees and attributes of nodes (FRG numbers, N, fitted values (in dollars)); flag cases where increasing FIM scores result in higher predicted payment. These may lead to politically unacceptable payment formulas.
- Because there tends to be a monotone decreasing relationship between cost and FIM scores, at least where effects are strong, non-monotonicities tend to occur at the bottom of trees. Joining adjacent bottom nodes of a tree can eliminate these. These occur most frequently when CART is inconsistent in its splits. Consider the case where we have a split on motor score and then the lower motor score group splits on cognitive function. The last split may introduce a discontinuity: The low motor/high cognitive group could cost less than the high motor group. This kind of non-monotonicity is an artifact of CART rather than the result of actual cost patterns. After CART produces its tree, we join the deepest adjacent nodes of the tree to eliminate situations where adding a point to any of the FIM components would result in a higher predicted cost. We perform this process recursively until the tree satisfies the required monotonicity constraints (the higher the FIM score, the lower the cost).

- Perform additional pruning on adjacent nodes where fitted values
  are close (i.e., $1,500). Repeat this step so long as adjacent
  nodes are within $1,500, but do not join any nodes that would
  result in a predicted value that differs from the original by more
  than $1,000.

Table 3.13 shows how the number of nodes varied at each stage of
this process. 1SD-CART started with 126 total nodes and expanded to 136
with the addition of nodes that boosted R-squared. The monotonicity
requirement further pared the total down to 118, and the pruning for close
cost outcomes reduced it to 95. Table 3.14 shows the aggregate standard
deviations for each step of this process. The monotonicity requirement and
subsequent pruning affected RMSEs minimally, less than 0.002. Table 3.15
displays the 136-node model prior to the pruning. The letters in column 4
show what was grouped to accommodate both monotonicity and proximity of
adjacent nodes. For example, grouping (d) was made for monotonicity
purposes. Grouping (f) joined the first four lines to remove monotonicity
violations; a fifth line was later added because of node value proximity.
Table 3.16 displays the final 95-node model, shown in a format similar to
that of Table 3.13 of the interim report.

Table 3.13

Number of Nodes at Various Stages of Pruning

RIC	1SD	ADD	MON	FNL	Description
01	18	18	18	14	Stroke
02	5	5	5	5	Brain dysfunction, traumatic
03	4	4	4	4	Brain dysfunction, nontraumatic
04	5	5	4	4	Spinal cord dysfunction, traumatic
05	6	6	6	5	Spinal cord dysfunction, nontraumatic
06	4	4	4	4	Neurological conditions
07	16	16	10	5	Orthopedic, lower extremity fracture
08	22	22	12	6	Orthopedic, lower extremity joint replacement
09	6	6	6	4	Orthopedic, other
10	3	5	5	5	Amputation, lower extremity
11	2	3	3	3	Amputation, other
12	4	6	6	5	Osteoarthritis
13	3	4	4	4	Rheumatoid and other arthritis
14	4	4	4	4	Cardiac
15	4	4	4	4	Pulmonary
16	4	4	3	2	Pain syndrome
17	3	3	3	3	Major multiple trauma, without injury to brain or spinal cord
18	2	4	4	4	Major multiple trauma, with injury to brain or spinal cord
19	2	3	3	3	Guillain-Barré syndrome
20	8	8	8	5	Other disabling impairments
21	1	2	2	2	Burns
Total	126	136	118	95	

Notes: 1SD = 1SD-CART; ADD = 1SD plus nodes that increase R-squared; MON = ADD after pruning for non-monotonicities; FNL = MON after pruning where exponentiated averages are "close."

Table 3.14

RMSEs at Various Stages of Pruning

Fit Year	Eval. Year	1SD	ADD	MON	FNL
99	96	0.474	0.474	0.474	0.475
	97	0.479	0.479	0.479	0.480
	98	0.483	0.483	0.483	0.485
	99	0.484	0.483	0.484	0.486

Notes: 1SD = 1SD-CART; ADD = 1SD plus nodes that increase R-squared; MON = ADD after pruning for non-monotonicities; FNL = MON after pruning where exponentiated averages are "close."

Table 3.15

**136-Node FRG Model, Before Correcting for Non-monotonicities
and Proximity**

RIC	FRG	N	Cost ($/year)	Grouping	Condition
01	18	4,215	20,869		M<41.5 & M<33.5 & A<81.5 & M<26.5
	17	3,763	18,233		M<41.5 & M<33.5 & A<81.5 & M>26.5
	16	1,065	18,546		M<41.5 & M<33.5 & A>81.5 & A<88.5 & M<26.5
	15	1,003	16,252		M<41.5 & M<33.5 & A>81.5 & A<88.5 & M>26.5
	14	584	14,750		M<41.5 & M<33.5 & A>81.5 & A>88.5
	13	3,620	15,756		M<41.5 & M>33.5 & M<38.5 & A<82.5
	12	987	13,739		M<41.5 & M>33.5 & M<38.5 & A>82.5
	11	3,102	13,616		M<41.5 & M>33.5 & M>38.5
	10	4,663	12,149	(a)	M>41.5 & M<52.5 & M<46.5 & C<31.5
	09	1,090	11,037	(a)	M>41.5 & M<52.5 & M<46.5 & C>31.5
	08	2,838	10,754	(b)	M>41.5 & M<52.5 & M>46.5 & C<28.5 & M<50.5
	07	1,254	9,779	(b)	M>41.5 & M<52.5 & M>46.5 & C<28.5 & M>50.5
	06	2,628	9,265	(b)	M>41.5 & M<52.5 & M>46.5 & C>28.5
	05	2,413	8,822	(c)	M>41.5 & M>52.5 & M<58.5 & C<29.5
	04	1,718	7,325	(c)	M>41.5 & M>52.5 & M<58.5 & C>29.5
	03	551	7,927		M>41.5 & M>52.5 & M>58.5 & C<22.5
	02	1,596	6,400		M>41.5 & M>52.5 & M>58.5 & C>22.5 & M<68.5
	01	250	5,064		M>41.5 & M>52.5 & M>58.5 & C>22.5 & M>68.5
02	05	428	19,149		M<39.5 & M<29.5
	04	400	14,101		M<39.5 & M>29.5
	03	602	11,522		M>39.5 & C<23.5
	02	303	9,858		M>39.5 & C>23.5 & M<51.5
	01	320	7,137		M>39.5 & C>23.5 & M>51.5
03	04	442	20,333		M<40.5 & M<24.5
	03	1,099	14,429		M<40.5 & M>24.5
	02	1,164	10,754		M>40.5 & M<50.5
	01	1,053	8,168		M>40.5 & M>50.5
04	05	111	14,913	(d)	M<35.5 & M<18.5 & A<55.5
	04	171	26,635	(d)	M<35.5 & M<18.5 & A>55.5
	03	604	16,236		M<35.5 & M>18.5
	02	599	11,282		M>35.5 & M<49.5
	01	398	7,785		M>35.5 & M>49.5
05	06	1,139	16,882		M<40.5 & M<33.5
	05	923	11,837		M<40.5 & M>33.5
	04	955	9,321	(e)	M>40.5 & M<50.5 & M<45.5
	03	1,079	8,063	(e)	M>40.5 & M<50.5 & M>45.5
	02	243	7,951		M>40.5 & M>50.5 & C<29.5
	01	1,498	6,317		M>40.5 & M>50.5 & C>29.5
06	04	2,221	13,373		M<46.5 & M<35.5

Table 3.15 (cont.)

RIC	FRG	N	Cost ($/year)	Grouping	Condition
	03	2,924	10,982		M<46.5 & M>35.5
	02	2,486	8,911		M>46.5 & M<55.5
	01	1,244	6,988		M>46.5 & M>55.5
07	16	374	11,861	(f)	M<45.5 & M<37.5 & M<33.5 & C<13.5
	15	2,330	13,602	(f)	M<45.5 & M<37.5 & M<33.5 & C>13.5 & C<33.5
	14	149	10,007	(f)	M<45.5 & M<37.5 & M<33.5 & C>13.5 & C>33.5 & C<34.5
	13	355	13,135	(f)	M<45.5 & M<37.5 & M<33.5 & C>13.5 & C>33.5 & C>34.5
	12	2,189	11,885	(f)	M<45.5 & M<37.5 & M>33.5
	11	1,471	11,603	(g)	M<45.5 & M>37.5 & M<41.5 & C<30.5
	10	1,562	10,583	(g)	M<45.5 & M>37.5 & M<41.5 & C>30.5
	09	706	10,394	(h)	M<45.5 & M>37.5 & M>41.5 & A<81.5 & C<30.5
	08	1,264	9,284	(h)	M<45.5 & M>37.5 & M>41.5 & A<81.5 & C>30.5
	07	1,597	10,530	(h)	M<45.5 & M>37.5 & M>41.5 & A>81.5
	06	1,995	9,274	(i)	M>45.5 & M<51.5 & C<31.5
	05	1,551	8,665	(i)	M>45.5 & M<51.5 & C>31.5 & M<48.5
	04	1,593	7,919	(i)	M>45.5 & M<51.5 & C>31.5 & M>48.5
	03	349	8,787	(j)	M>45.5 & M>51.5 & M<55.5 & C<29.5
	02	1,792	7,245	(j)	M>45.5 & M>51.5 & M<55.5 & C>29.5
	01	1,350	6,380	(j)	M>45.5 & M>51.5 & M>55.5
08	22	1,411	11,237	(k)	M<46.5 & C<31.5 & M<39.5 & M<34.5
	21	1,323	10,007	(k)	M<46.5 & C<31.5 & M<39.5 & M>34.5
	20	3,083	8,544		M<46.5 & C<31.5 & M>39.5
	19	749	8,647	(l)	M<46.5 & C>31.5 & M<42.5 & A<80.5 & C<34.5 & C<33.5
	18	408	5,603	(l)	M<46.5 & C>31.5 & M<42.5 & A<80.5 & C<34.5 & C>33.5 & M<36.5
	17	404	7,879	(l)	M<46.5 & C>31.5 & M<42.5 & A<80.5 & C<34.5 & C>33.5 & M>36.5
	16	499	9,576	(l)	M<46.5 & C>31.5 & M<42.5 & A<80.5 & C>34.5 & M<35.5
	15	1,607	7,927	(l)	M<46.5 & C>31.5 & M<42.5 & A<80.5 & C>34.5 & M>35.5
	14	1,095	9,100	(l)	M<46.5 & C>31.5 & M<42.5 & A>80.5
	13	2,153	6,898	(l)	M<46.5 & C>31.5 & M>42.5 & A<74.5
	12	2,408	7,518	(l)	M<46.5 & C>31.5 & M>42.5 & A>74.5
	11	2,674	7,023	(m)	M>46.5 & M<54.5 & C<33.5 & A<81.5 & M<50.5
	10	2,259	6,400	(m)	M>46.5 & M<54.5 & C<33.5 & A<81.5 & M>50.5
	09	1,378	7,548	(m)	M>46.5 & M<54.5 & C<33.5 & A>81.5
	08	3,594	6,438	(m)	M>46.5 & M<54.5 & C>33.5 & M<49.5 & A<81.5

Table 3.15 (cont.)

RIC	FRG	N	Cost ($/year)	Grouping	Condition
	07	623	7,274	(m)	M>46.5 & M<54.5 & C>33.5 & M<49.5 & A>81.5
	06	5,406	5,779	(m)	M>46.5 & M<54.5 & C>33.5 & M>49.5 & A<77.5
	05	2,353	6,260	(m)	M>46.5 & M<54.5 & C>33.5 & M>49.5 & A>77.5
	04	653	6,342	(n)	M>46.5 & M>54.5 & M<57.5 & C<31.5
	03	4,186	5,443	(n)	M>46.5 & M>54.5 & M<57.5 & C>31.5
	02	3,839	5,162	(o)	M>46.5 & M>54.5 & M>57.5 & M<62.5
	01	1,322	4,666	(o)	M>46.5 & M>54.5 & M>57.5 & M>62.5
09	06	1,777	11,920		M<46.5 & M<37.5
	05	1,202	10,148	(p)	M<46.5 & M>37.5 & C<30.5
	04	1,713	8,822	(p)	M<46.5 & M>37.5 & C>30.5
	03	2,905	7,692		M>46.5 & M<53.5
	02	432	6,857	(q)	M>46.5 & M>53.5 & C<31.5
	01	1,281	5,722	(q)	M>46.5 & M>53.5 & C>31.5
10	05	1,495	14,794		M<45.5 & M<38.5
	04	1,319	12,531		M<45.5 & M>38.5
	03	1,465	10,938		M>45.5 & M<51.5
	02	1,412	9,377		M>45.5 & M>51.5 & M<60.5
	01	465	7,809		M>45.5 & M>51.5 & M>60.5
11	03	217	14,300		M<51.5 & M<37.5
	02	580	10,711		M<51.5 & M>37.5
	01	407	7,793		M>51.5
12	06	861	12,772		M<47.5 & M<38.5
	05	1,403	10,342		M<47.5 & M>38.5
	04	862	8,920	(r)	M>47.5 & M<54.5 & C<33.5
	03	781	7,785	(r)	M>47.5 & M<54.5 & C>33.5
	02	489	7,578		M>47.5 & M>54.5 & C<33.5
	01	640	6,027		M>47.5 & M>54.5 & C>33.5
13	04	397	13,427		M<46.5 & M<35.5
	03	710	10,097		M<46.5 & M>35.5
	02	660	8,358		M>46.5 & M<53.5
	01	583	6,667		M>46.5 & M>53.5
14	04	1,018	12,657		M<47.5 & M<37.5
	03	2,200	9,887		M<47.5 & M>37.5
	02	2,747	7,871		M>47.5 & M<55.5
	01	2,139	6,298		M>47.5 & M>55.5
15	04	629	15,460		M<47.5 & M<35.5
	03	1,367	11,328		M<47.5 & M>35.5
	02	2,461	9,072		M>47.5 & M<60.5
	01	925	7,662		M>47.5 & M>60.5
16	04	958	9,789		M<44.5
	03	981	7,856	(s)	M>44.5 & M<52.5
	02	166	8,656	(s)	M>44.5 & M>52.5 & A<63.5
	01	888	6,323	(s)	M>44.5 & M>52.5 & A>63.5
17	03	333	15,139		M<45.5 & M<32.5
	02	750	11,339		M<45.5 & M>32.5

Table 3.15 (cont.)

RIC	FRG	N	Cost ($/year)	Grouping	Condition
	01	596	8,136		M>45.5
18	04	87	25,336		M<44.5 & M<25.5
	03	229	14,516		M<44.5 & M>25.5
	02	103	9,927		M>44.5 & C<32.5
	01	58	6,470		M>44.5 & C>32.5
19	03	143	25,591		M<46.5 & M<30.5
	02	245	16,916		M<46.5 & M>30.5
	01	224	9,274		M>46.5
20	08	1,812	14,415		M<44.5 & M<32.5 & A<81.5
	07	854	12,173		M<44.5 & M<32.5 & A>81.5
	06	2,134	11,861	(t)	M<44.5 & M>32.5 & M<37.5
	05	4,405	10,530	(t)	M<44.5 & M>32.5 & M>37.5
	04	4,305	9,145	(u)	M>44.5 & M<53.5 & M<49.5
	03	3,322	8,259	(u)	M>44.5 & M<53.5 & M>49.5
	02	3,132	7,347	(v)	M>44.5 & M>53.5 & M<59.5
	01	1,589	6,400	(v)	M>44.5 & M>53.5 & M>59.5
21	02	119	17,677		M<45.5
	01	94	9,284		M>45.5

Notes: M stands for the 12-component FIM motor score, C for the standard FIM cognitive score, and A for age. The FRG numbers were assigned by CART mostly in increasing order of average cost, although exceptions were made to keep adjacent nodes together.

Table 3.16

Recommended 95-Node FRG Model

RIC	FRG	N	Cost ($/year)	Condition
01	14	4,215	20,869	M<41.5 & M<33.5 & A<81.5 & M<26.5
	13	3,763	18,233	M<41.5 & M<33.5 & A<81.5 & M>26.5
	12	1,065	18,546	M<41.5 & M<33.5 & A>81.5 & A<88.5 & M<26.5
	11	1,003	16,252	M<41.5 & M<33.5 & A>81.5 & A<88.5 & M>26.5
	10	584	14,750	M<41.5 & M<33.5 & A>81.5 & A>88.5
	09	3,620	15,756	M<41.5 & M>33.5 & M<38.5 & A<82.5
	08	987	13,739	M<41.5 & M>33.5 & M<38.5 & A>82.5
	07	3,102	13,616	M<41.5 & M>33.5 & M>38.5
	06	5,753	11,932	M>41.5 & M<52.5 & M<46.5
	05	6,720	9,967	M>41.5 & M<52.5 & M>46.5
	04	4,131	8,168	M>41.5 & M>52.5 & M<58.5
	03	551	7,927	M>41.5 & M>52.5 & M>58.5 & C<22.5
	02	1,596	6,400	M>41.5 & M>52.5 & M>58.5 & C>22.5 & M<68.5
	01	250	5,064	M>41.5 & M>52.5 & M>58.5 & C>22.5 & M>68.5
02	05	428	19,149	M<39.5 & M<29.5
	04	400	14,101	M<39.5 & M>29.5
	03	602	11,522	M>39.5 & C<23.5
	02	303	9,858	M>39.5 & C>23.5 & M<51.5
	01	320	7,137	M>39.5 & C>23.5 & M>51.5
03	04	442	20,333	M<40.5 & M<24.5
	03	1,099	14,429	M<40.5 & M>24.5
	02	1,164	10,754	M>40.5 & M<50.5
	01	1,053	8,168	M>40.5 & M>50.5
04	04	282	21,248	M<35.5 & M<18.5
	03	604	16,236	M<35.5 & M>18.5
	02	599	11,282	M>35.5 & M<49.5
	01	398	7,785	M>35.5 & M>49.5
05	05	1,139	16,882	M<40.5 & M<33.5
	04	923	11,837	M<40.5 & M>33.5
	03	2,034	8,630	M>40.5 & M<50.5
	02	243	7,951	M>40.5 & M>50.5 & C<29.5
	01	1,498	6,317	M>40.5 & M>50.5 & C>29.5
06	04	2,221	13,373	M<46.5 & M<35.5
	03	2,924	10,982	M<46.5 & M>35.5
	02	2,486	8,911	M>46.5 & M<55.5
	01	1,244	6,988	M>46.5 & M>55.5
07	05	5,397	12,620	M<45.5 & M<37.5
	04	3,033	11,059	M<45.5 & M>37.5 & M<41.5
	03	3,567	10,047	M<45.5 & M>37.5 & M>41.5
	02	5,139	8,656	M>45.5 & M<51.5
	01	3,491	7,030	M>45.5 & M>51.5
08	06	2,734	10,625	M<46.5 & C<31.5 & M<39.5
	05	3,083	8,544	M<46.5 & C<31.5 & M>39.5
	04	9,323	7,708	M<46.5 & C>31.5

Table 3.16 (cont.)

RIC	FRG	N	Cost ($/year)	Condition
	03	18,287	6,393	M>46.5 & M<54.5
	02	4,839	5,552	M>46.5 & M>54.5 & M<57.5
	01	5,161	5,029	M>46.5 & M>54.5 & M>57.5
09	04	1,777	11,920	M<46.5 & M<37.5
	03	2,915	9,339	M<46.5 & M>37.5
	02	2,905	7,692	M>46.5 & M<53.5
	01	1,713	5,991	M>46.5 & M>53.5
10	05	1,495	14,794	M<45.5 & M<38.5
	04	1,319	12,531	M<45.5 & M>38.5
	03	1,465	10,938	M>45.5 & M<51.5
	02	1,412	9,377	M>45.5 & M>51.5 & M<60.5
	01	465	7,809	M>45.5 & M>51.5 & M>60.5
11	03	217	14,300	M<51.5 & M<37.5
	02	580	10,711	M<51.5 & M>37.5
	01	407	7,793	M>51.5
12	05	861	12,772	M<47.5 & M<38.5
	04	1,403	10,342	M<47.5 & M>38.5
	03	1,643	8,358	M>47.5 & M<54.5
	02	489	7,578	M>47.5 & M>54.5 & C<33.5
	01	640	6,027	M>47.5 & M>54.5 & C>33.5
13	04	397	13,427	M<46.5 & M<35.5
	03	710	10,097	M<46.5 & M>35.5
	02	660	8,358	M>46.5 & M<53.5
	01	583	6,667	M>46.5 & M>53.5
14	04	1,018	12,657	M<47.5 & M<37.5
	03	2,200	9,887	M<47.5 & M>37.5
	02	2,747	7,871	M>47.5 & M<55.5
	01	2,139	6,298	M>47.5 & M>55.5
15	04	629	15,460	M<47.5 & M<35.5
	03	1,367	11,328	M<47.5 & M>35.5
	02	2,461	9,072	M>47.5 & M<60.5
	01	925	7,662	M>47.5 & M>60.5
16	02	958	9,789	M<44.5
	01	2,035	7,201	M>44.5
17	03	333	15,139	M<45.5 & M<32.5
	02	750	11,339	M<45.5 & M>32.5
	01	596	8,136	M>45.5
18	04	87	25,336	M<44.5 & M<25.5
	03	229	14,516	M<44.5 & M>25.5
	02	103	9,927	M>44.5 & C<32.5
	01	58	6,470	M>44.5 & C>32.5
19	03	143	25,591	M<46.5 & M<30.5
	02	245	16,916	M<46.5 & M>30.5
	01	224	9,274	M>46.5

Table 3.16 (cont.)

RIC	FRG	N	Cost ($/year)	Condition
20	05	1,812	14,415	M<44.5 & M<32.5 & A<81.5
	04	854	12,173	M<44.5 & M<32.5 & A>81.5
	03	6,539	10,949	M<44.5 & M>32.5
	02	7,627	8,752	M>44.5 & M<53.5
	01	4,721	7,016	M>44.5 & M>53.5
21	02	119	17,677	M<45.5
	01	94	9,284	M>45.5

Notes: M stands for the 12-component FIM motor score, C for the standard FIM cognitive score, and A for age. The FRG numbers were assigned by CART mostly in increasing order of average cost, although exceptions were made to keep adjacent nodes together.

Performance of Recommended FRGs

The recommended FRGs will be used in a payment formula that adjusts its predicted costs by a variety of factors: hospital characteristics, comorbidities, budget neutrality requirements, etc. CART provided both classification groups and predictions of the average log-adjusted cost within groups. But CART is a constrained fitting procedure. We wanted to know whether a more complex formula, one that adheres to fewer constraints and has the ability to trace complex cost curves, would pay hospitals differently in total.

To answer this question, we performed a simulation exercise utilizing the recommended FRGs and two MART models described previously. The first MART model used the same index (M12C5) used in the FRGs. The second MART model used the gold standard index (StJe5). For each MART model, we asked what the total payments to hospitals would be if we used the models' prediction formula rather than the FRG payment formula.

Using the two different versions of MART answers two types of questions. First, given that nobody really contemplates going beyond a motor and cognitive index at this point, the MART-MC12C5 results tell us how sensitive total payments are to model functional forms. The payment system will be fair if these differences are small. On the other hand, the MART-StJe5 results tell us whether there is useful predictive information in a five-component set of indices. We saw some evidence for this in the RMSE computations shown in the previous subsection. If these effects

persist at the aggregate level, we ultimately want to know why and whether there is a deficiency in FRGs that can be eliminated.

The dependent variable fit in all these models is

$$log(TCOST / [wage\_adjustment]) .$$

We transformed fitted values back to dollars by

$$[prediction] = exp(fit) * [wage\_adjustment] * constant ,$$

where the constant is chosen to assure budget neutrality (i.e., it makes predictions sum to costs). We then aggregated the predicted values by hospital for each year, and we computed payment (to payment) ratios within hospitals under alternative model pairs. Use of this ratio cancels out facility-level payment adjustments (Section 7), so this becomes a true comparison of payments assuming no adjustment for comorbidities.

Table 3.17 summarizes the results for the MART-M12C5 versus FRG payment formulas in each of four years. For each prediction year, over 90 percent of hospitals have payment ratios between 0.98 and 1.02, and only about 1 percent of hospitals have payment ratios outside of 0.96 to 1.04. In case-weighted comparisons, about 94 percent of the hospitals are between 0.98 and 1.02. This percentage is remarkably close to 100 percent and affirms that the CART prediction formula would pay about the same in total as the more flexible MART prediction formula at the hospital level. We think this is strong confirmation for using our proposed set of FRGs.

The corresponding results from using MART-StJe5 are more equivocal. Table 3.18 shows that the distribution of hospital payment ratios is more spread out. Only about 60 percent fall between 0.98 and 1.02, and about 13 percent are outside of 0.96 to 1.04. The case-weighted numbers are a little closer to 100 percent but still considerably more dispersed than the MART-M12C5 numbers.

Table 3.17

Hospital Payment Ratios for Final FRGs Versus MART--M12C5

Hospital Payment Ratio	Percentage of Hospitals (facility-weighted)				Percentage of Hospitals (case-weighted)			
	1996	1997	1998	1999	1996	1997	1998	1999
0.90	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.94	0.2	0.0	0.0	0.0	0.1	0.0	0.0	0.0
0.95	0.0	0.2	0.4	0.1	0.0	0.0	0.2	0.1
0.96	0.9	0.8	1.0	1.0	0.4	0.3	0.3	1.1
0.97	2.5	3.5	2.5	2.9	2.5	2.5	2.4	2.1
0.98	10.7	10.0	10.6	9.5	11.6	9.8	11.9	8.0
0.99	20.7	22.7	20.8	21.6	21.6	25.8	21.8	24.6
1.00	27.4	26.9	27.9	26.0	28.9	28.4	30.9	29.7
1.01	21.2	21.5	22.2	23.9	22.3	22.0	21.5	24.3
1.02	11.6	10.5	9.6	10.4	9.7	8.3	8.3	7.0
1.03	2.2	3.2	2.4	3.5	1.9	2.7	2.2	2.4
1.04	1.6	0.5	1.5	0.9	0.8	0.2	0.5	0.6
1.05	0.4	0.2	0.4	0.3	0.1	0.1	0.1	0.2
1.06	0.4	0.0	0.3	0.0	0.2	0.0	0.0	0.0
1.07	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

At this point, we do not know the reasons why the FRGs and MART-StJe5 diverge in their prediction of total hospital payments. Given that the FRGs and MART-M12C5 agree so well, we think that MART-StJe5 may be producing cost curves that are not monotone and hence are not compatible with the design criteria for a payment system. This finding was of little importance in the RMSE investigation because it led to only slight underestimates of percentage of variance explained, but it matters much more here. In any case, we plan to investigate reasons for this discrepancy in phase II of our work (see Section 10).

Table 3.18

Hospital Payment Ratios for Final FRGs Versus MART-StJe5

Hospital Payment Ratio	Percentage of Hospitals (facility-weighted)				Percentage of Hospitals (case-weighted)			
	1996	1997	1998	1999	1996	1997	1998	1999
0.83	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.89	0.2	0.0	0.0	0.1	0.0	0.0	0.0	0.2
0.91	0.2	0.2	0.3	0.1	0.2	0.2	0.2	0.1
0.92	0.7	0.8	0.4	0.3	0.3	0.4	0.8	0.4
0.93	0.7	1.0	0.4	0.7	0.6	0.3	0.5	0.4
0.94	1.3	1.8	1.5	1.9	1.1	1.3	1.1	2.0
0.95	3.1	2.7	2.4	2.6	3.1	2.0	1.5	2.5
0.96	6.9	6.3	4.8	4.7	5.9	5.4	4.4	3.7
0.97	5.6	7.8	7.6	7.9	6.1	8.5	7.5	8.6
0.98	12.3	11.0	10.2	8.1	13.7	12.7	10.7	9.1
0.99	12.3	12.9	14.7	15.0	11.7	13.3	16.2	16.3
1.00	14.3	12.9	14.1	12.1	16.4	15.0	15.8	12.6
1.01	10.9	11.4	12.7	15.3	11.2	12.6	13.0	16.2
1.02	10.3	10.7	9.7	9.6	9.6	11.0	10.3	9.7
1.03	8.0	7.3	9.1	9.5	9.0	6.7	8.9	9.1
1.04	6.3	6.0	5.1	4.5	6.4	5.7	3.9	3.5
1.05	3.1	2.7	3.0	3.3	2.0	2.4	3.2	2.6
1.06	1.8	2.7	1.6	2.6	1.8	1.4	1.1	2.0
1.07	0.9	0.7	0.9	0.4	0.5	0.5	0.3	0.3
1.08	0.4	0.2	0.4	0.6	0.2	0.1	0.2	0.3
1.09	0.2	0.7	0.1	0.4	0.1	0.2	0.1	0.4
1.10	0.2	0.3	0.3	0.3	0.1	0.3	0.1	0.1
1.11	0.2	0.0	0.1	0.0	0.0	0.0	0.0	0.0
1.12	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Year-to-Year Stability of FRGs

The final trees differ in some respects from the trees produced in the interim report. This is not surprising. CART is trying to fit step functions to continuous curves, so the cut-points are imprecisely determined. Further, we now use two years of data for small RICs, thus increasing the number of FRGs in these areas. We think the important question is not whether the trees are identical but whether the tree models produce a consistent and accurate set of predictions. For now, we simply ask the question, If one used the current FRGs and associated

predictions, how different would these predictions be over the different years?

Figures 3.18-3.38 demonstrate that this set of FRG models fits the data pretty well in all years. They show the predicted means within FRGs; predictions are normalized to have the same mean as actual log-cost across all RICs. Except for the RICs that were pooled, the 1999 predicted means fall right on top of the log-cost averages (rightmost panels). As one moves to the left, one sees the models projected backwards in time. The predictions for 1998 look quite good. They are a little worse for 1997, and a little worse still for 1996.

As the system is refined over time, rules will need to be developed that determine whether the models should be refit and how this refit might occur. We will gain some insight into the performance of various algorithms--such as adjusting cut-points, splitting FRGs, combining FRGs, or completely refitting the CART models--by using these plots to identify anomalous patterns and seeing what it takes to eliminate them. Other issues that we will address in refinement include the value of using information from the new IRF patient assessment instrument (PAI) items as part of the classification and integrating information about comorbidities with the FRG analysis.
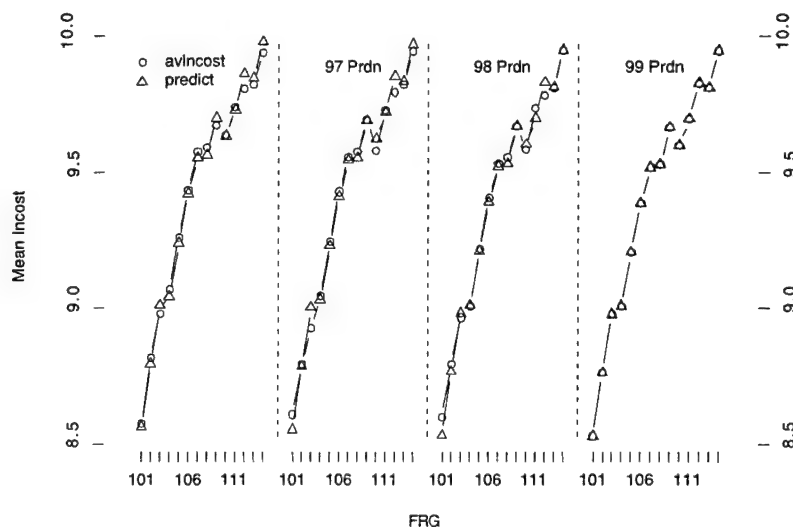
Figure 3.18--Actual and Predicted FRG Means: RIC=01, Fityear=99
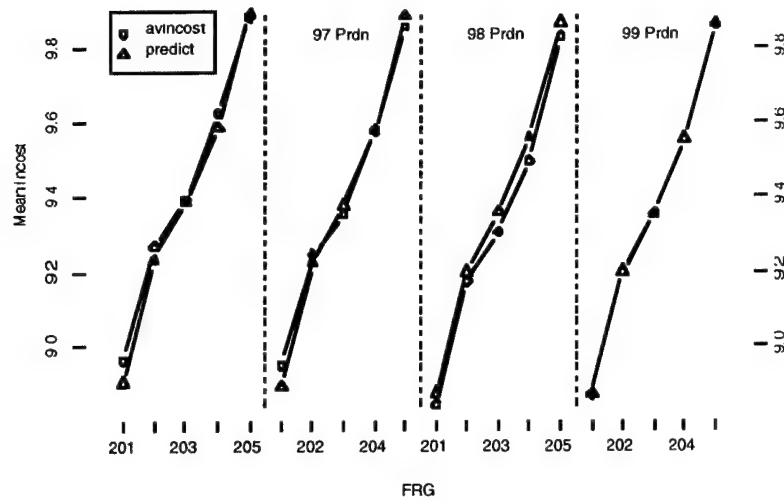
Figure 3.19--Actual and Predicted FRG Means: RIC=02, Fityear=99
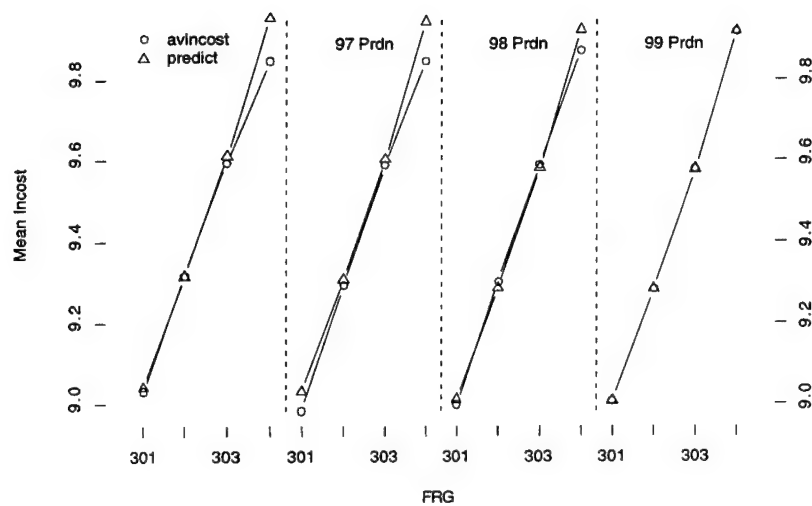
Figure 3.20--Actual and Predicted FRG Means: RIC=03, Fityear=99

- 87 -
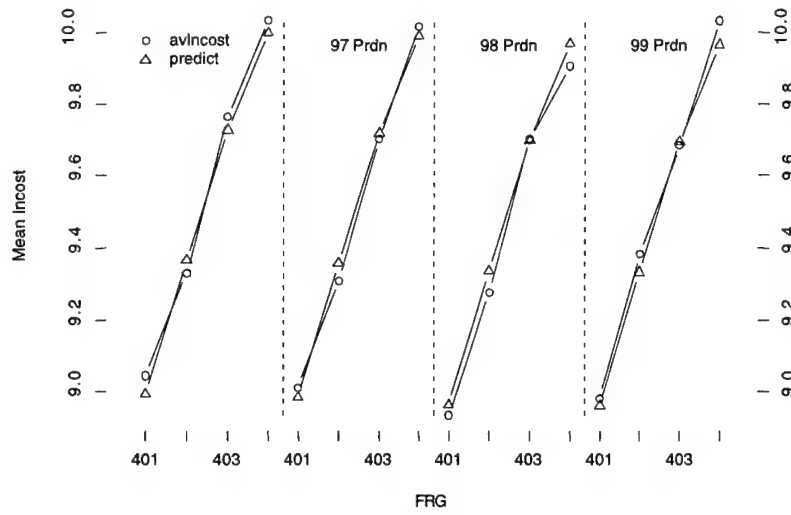
Figure 3.21--Actual and Predicted FRG Means: RIC=04, Fityear=98,99
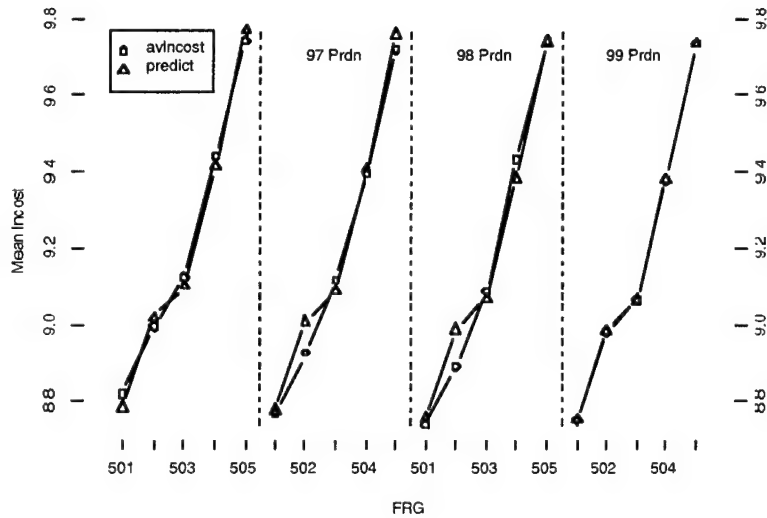
Figure 3.22--Actual and Predicted FRG Means: RIC=05, Fityear=99

- 88 -
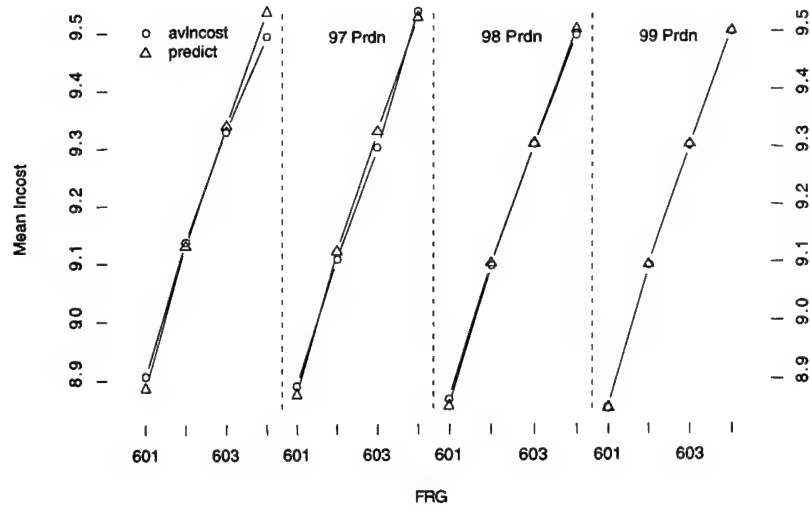
Figure 3.23--Actual and Predicted FRG Means: RIC=06, Fityear=99
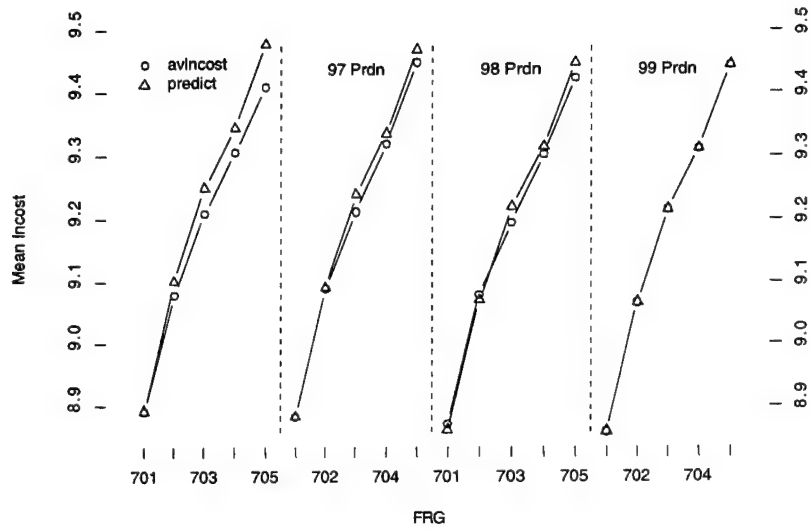
Figure 3.24--Actual and Predicted FRG Means: RIC=07, Fityear=99
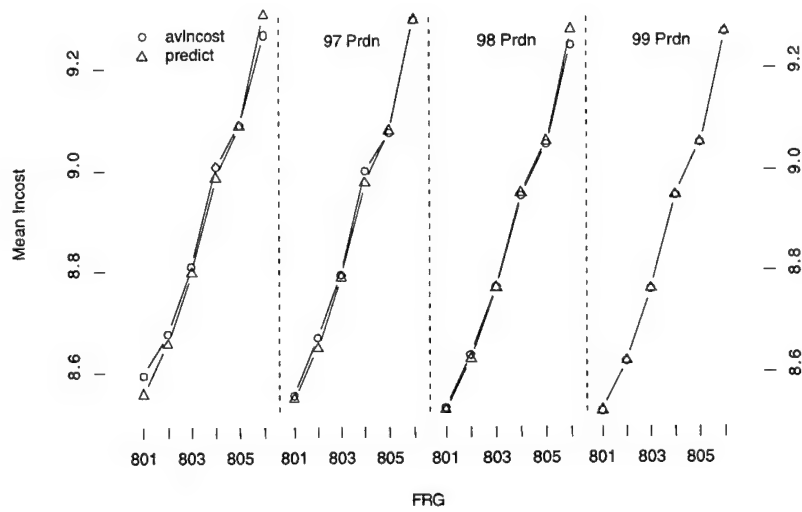
Figure 3.25--Actual and Predicted FRG Means: RIC=08, Fityear=99
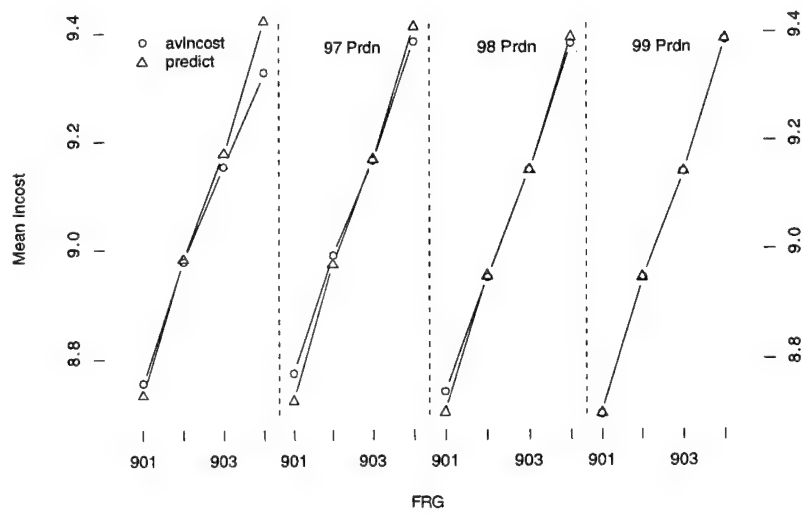
Figure 3.26--Actual and Predicted FRG Means: RIC=09, Fityear=99
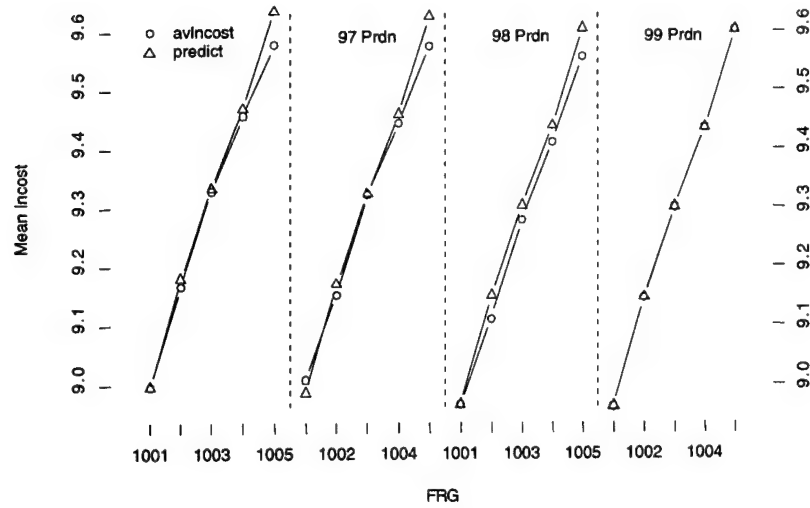
Figure 3.27--Actual and Predicted FRG Means: RIC=10, Fityear=99
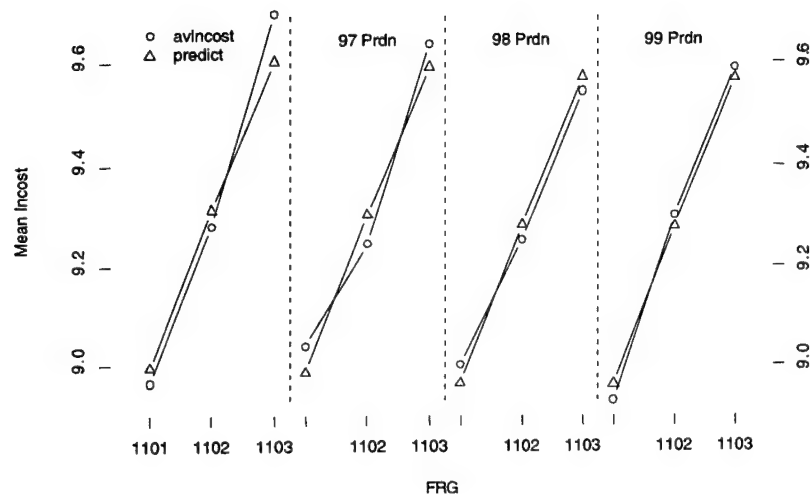
Figure 3.28--Actual and Predicted FRG Means: RIC=11, Fityear=98,99
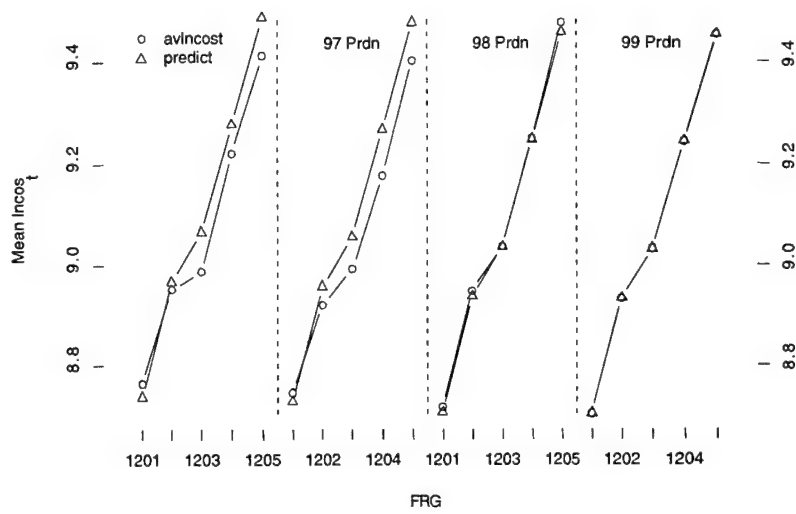
Figure 3.29--Actual and Predicted FRG Means: RIC=12, Fityear=99
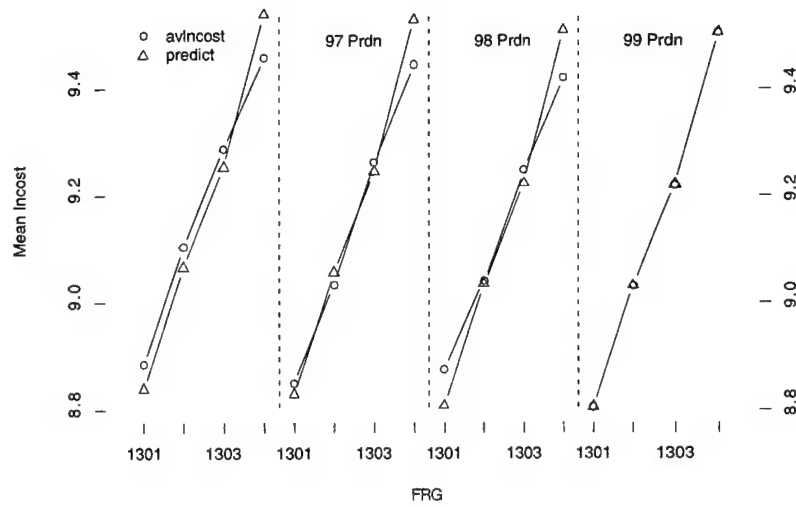
Figure 3.30--Actual and Predicted FRG Means: RIC=13, Fityear=99
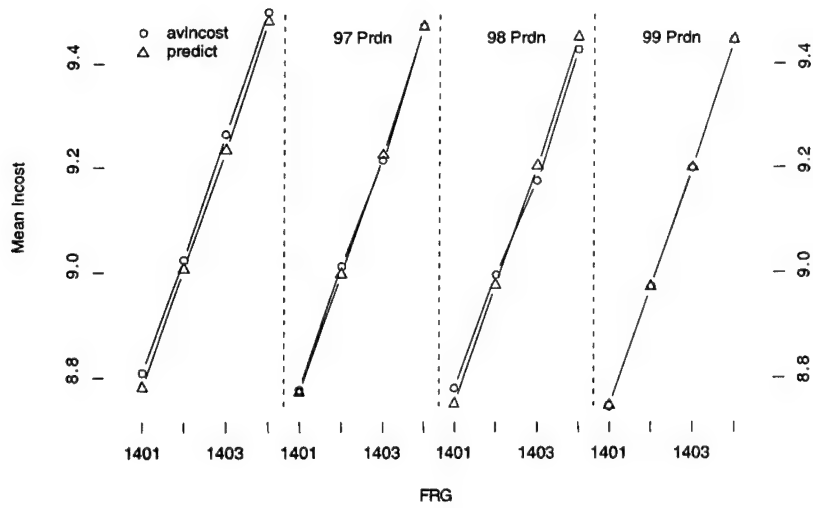
Figure 3.31--Actual and Predicted FRG Means: RIC=14, Fityear=99
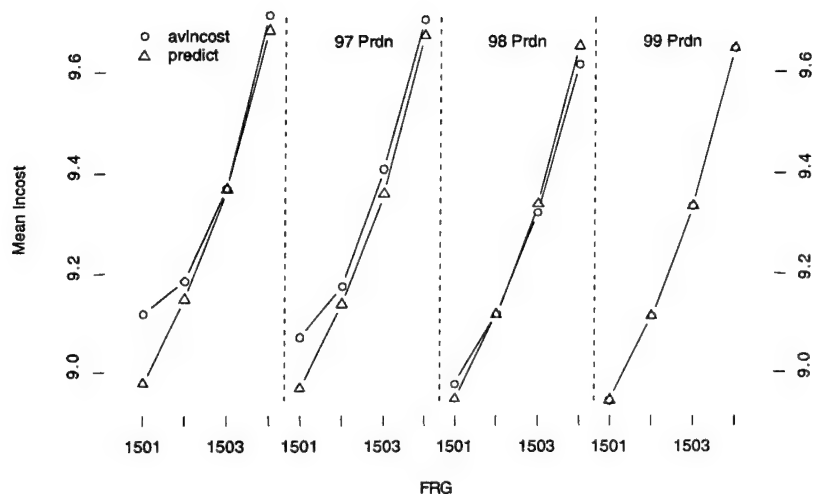
Figure 3.32--Actual and Predicted FRG Means: RIC=15, Fityear=99
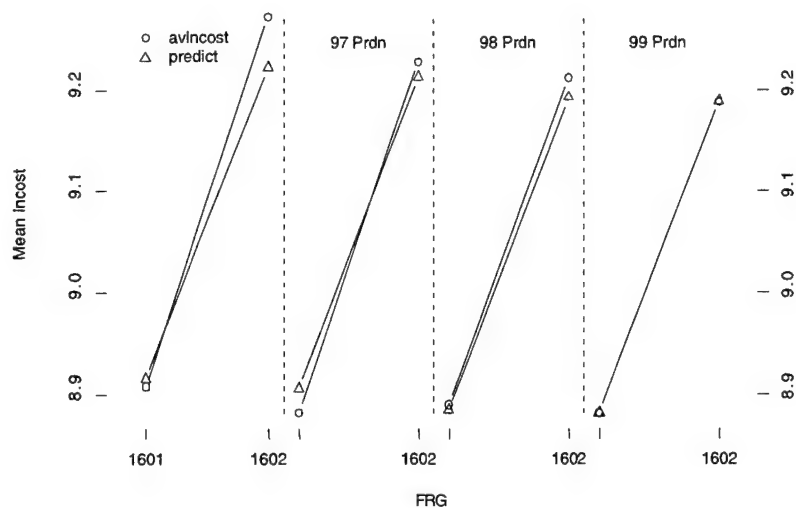
Figure 3.33--Actual and Predicted FRG Means: RIC=16, Fityear=99
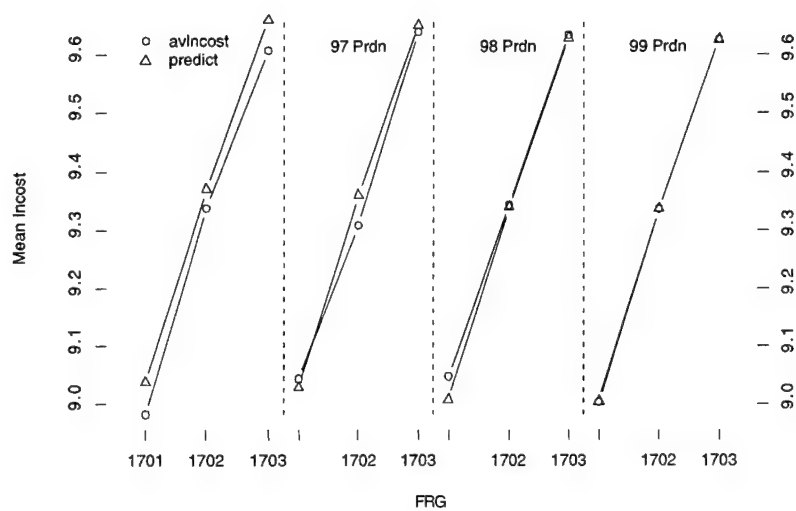
Figure 3.34--Actual and Predicted FRG Means: RIC=17, Fityear=99
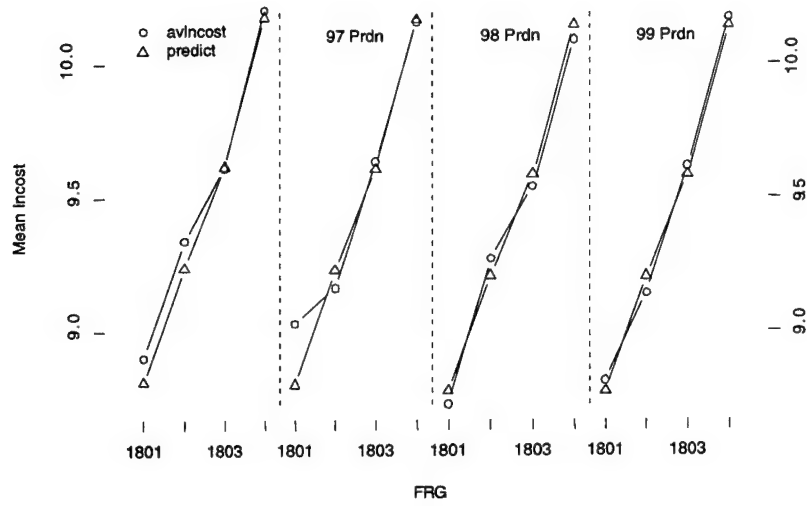
Figure 3.35--Actual and Predicted FRG Means: RIC=18, Fityear=98,99
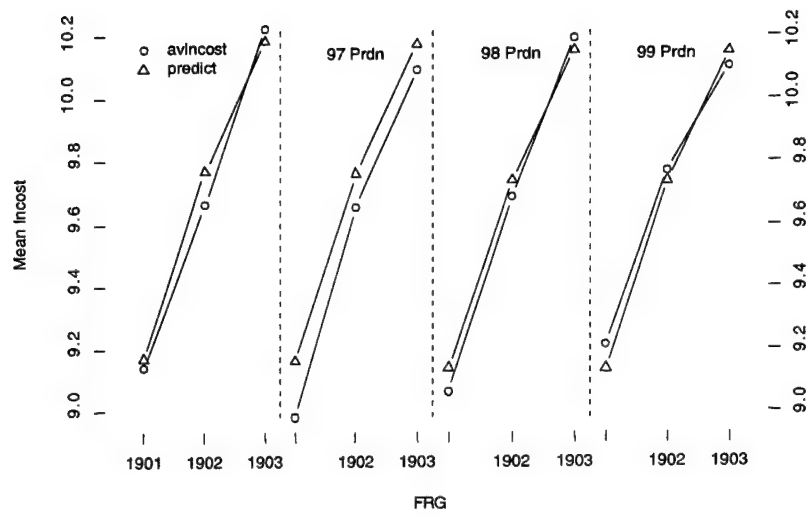
Figure 3.36--Actual and Predicted FRG Means: RIC=19, Fityear=98,99
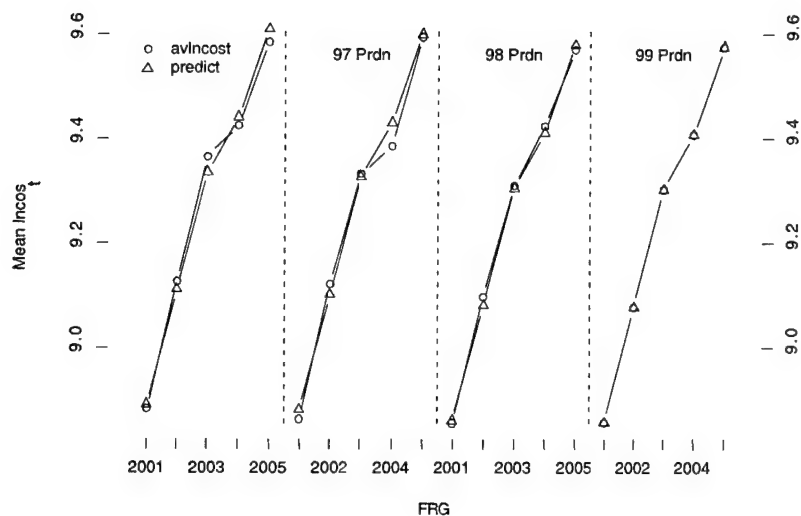
Figure 3.37--Actual and Predicted FRG Means: RIC=20, Fityear=99
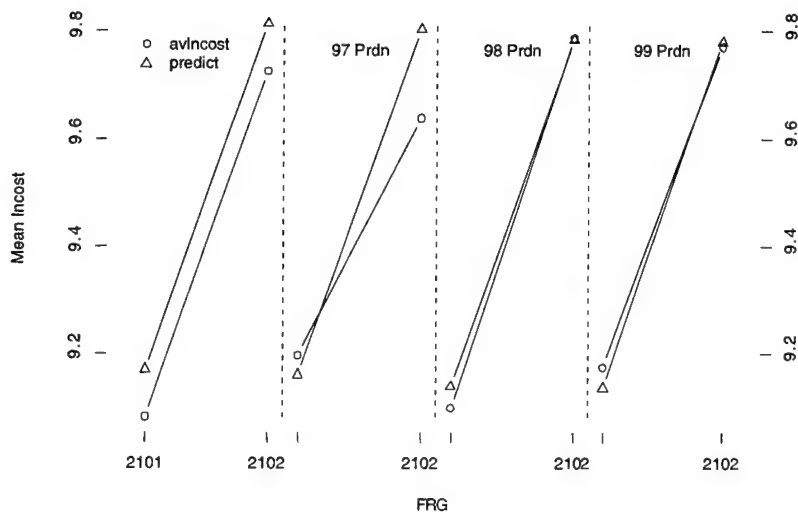
Figure 3.38--Actual and Predicted FRG Means: RIC=21, Fityear=98,99

CONCLUSIONS

We found that FRGs built using CART on the FIM data can predict cost reasonably well. They offer a parsimonious description of the payment formula while maintaining prediction accuracy. In addition, the payment formulas remained stable over four years, continuing to accurately predict cost.

CART produced a classification system that predicts cost with reasonable accuracy, has easily implemented payment rules, and avoids inconsistencies in the payments. While we are aware that CART has some limitations, we were able to compare its cost predictions with those produced by MART, a state-of-the-art flexible fitting method. We observed two key facts: CART explains more than 80 percent of the "explainable" variance in cost, and a simple payment formula that uses CART predictions will not pay hospitals much differently from MART.

We evaluated the performance of a number of indices against the standard 13-element motor score and 5-element cognitive score. We found that in the different RICs, both transfer to tub and comprehension often have the wrong sign in OLS regressions. That is, costs were higher when the functional independence measure was higher. Also, neither seemed to add to the explanatory power of cost prediction models. Because we believe that the UDSmr transfer-to-tub question, as currently worded, does not measure an absolute level of function, we recommended to CMS that this item not be used in creating FRGs. We could not use a similar rationale to argue for removing the comprehension item. We recognized that if the comprehension item were taken out of the index, the system would provide no extra incentives to treat patients with lowered comprehension. Therefore, we recommended leaving comprehension in the index.

We considered index sets that broke motor scores down into finer groups, such as separating ADLs from mobility measures, but chose not to use these sets when we observed that CART was unable to produce trees with a manageable number of nodes. We believe that asking CART to work with one or two indices in each RIC, rather than four or five, is a better way to explore alternative index sets. Phase II of this project will include a refinement phase during which we plan to examine a two-stage form of index selection: First select a couple of indices in each RIC using an alternative method (OLS and GAM, for example) and then have CART estimate cost from RIC-specific indices.

We observed that FRG models based on data from the most recent years fit well in all years. We compared actual and predicted FRG means for nonfitting years and saw that actual and predicted mean log-costs tended to be quite close. Also, the actual means in adjacent FRGs seemed to be

well separated from one another, indicating that the FRGs continued to discriminate between groups with differing levels of cost.

4. COMORBIDITIES

In this section, we discuss adding comorbidities and complications to the classification system. We began with a variety of hypotheses about which conditions affect cost, each suggested by a clinician. We tested each hypothesis to determine whether the nominated conditions were in fact associated with increased cost after controlling for FRG. These tests resulted in a list of conditions that are correlated with higher costs and that our clinical consultants believe actually cause an increase in costs. Finally, we evaluated alternative models for accounting for the cost of comorbidities. Some codes that were found to affect cost were eliminated from some payment models based on the judgment that they may be preventable.

DATA AND METHODS

Data

The analyses presented in this section were performed using data similar to those used to develop the FRG structure. Our basic analysis sample was reduced by excluding transfers, deaths, and cases with a length of stay of three days or less. We also eliminated as statistical outliers all cases that were beyond three standard deviations of the FRG mean in the log of cost. The FRG analysis used a slightly different definition of discharge home (see Section 2). In addition, the FRG analysis used the distribution of costs within RIC instead of FRG for the statistical outlier criterion. Thus, this analysis is based on a slightly different number of cases (1.4 percent more) than that in the last section. Table 4.1 shows the distribution of cases by RIC.

To identify patients with specific conditions, we used all the available secondary diagnoses. In particular, we included
- diagnoses 2 through 10 on the MEDPAR
- the seven fields found in items 30, 31, or 32 of the UDSmr FIM instrument (i.e., all ICD-9-CM diagnoses except the etiologic diagnosis)
- the secondary diagnosis on the HealthSouth data. (Although there was a field for secondary diagnosis on the COS data, it was never filled in.)

Most comorbidities are found on MEDPAR. In preliminary analyses we found that comorbidities found only on FIM had effects indistinguishable from those found on both FIM and MEDPAR or only on MEDPAR.[1]

Table 4.1

Sample Sizes for Comorbidity Analyses, by RIC

Rehabilitation Impairment Category	1996	1997	1998	1999
1 Stroke	33,212	36,399	37,459	37,864
2 Traumatic brain injury (BI)	1,414	1,15	1,910	2,097
3 Nontraumatic brain injury	2,581	2,981	3,451	3,823
4 Traumatic spinal cord injury (SCI)	748	849	946	981
5 Nontraumatic spinal cord injury	3,853	4,538	5,370	5,904
6 Neurological	4,779	5,931	7,945	9,028
7 Hip fracture	16,283	17,819	19,060	20,966
8 Replacement of lower extremity (LE) joint	31,652	38,409	41,396	43,936
9 Other orthopedic	5,352	6,719	8,171	9,516
10 Amputation, LE	4,887	5,636	5,989	6,227
11 Amputation, other	361	495	558	673
12 Osteoarthritis	2,417	2,958	4,015	5,105
13 Rheumatoid, other arthritis	1,179	1,572	1,966	2,392
14 Cardiac	4,158	5,772	6,977	8,220
15 Pulmonary	2,470	3,665	4,392	5,461
16 Pain syndrome	1,339	1,937	2,571	3,027
17 Major multiple trauma (MMT); no BI or SCI	1,205	1,355	1,567	1,704
18 MMT, with BI or SCI	164	226	227	265
19 Guillain-Barré	244	293	302	316
20 Miscellaneous	10,219	13,943	17,752	21,913
21 Burns	74	105	112	105
Total	128,591	153,317	172,136	189,523

NOTE: We use the definition of FRG found in Table 3.16.

The use of the FIM diagnoses introduces a complication: The UDSmr instrument allows up to seven codes, but the HealthSouth cases have only a single secondary diagnosis from the FIM instrument and the COS cases have none.[2] If it is true that the coding of comorbidities on the MEDPAR is similar for cases from each data source, then there should be more cases with missing

---

[1] In the preliminary analyses, we examined--separately for 1998 and 1999--the effect of adding a variable that was equal to 1 when the condition was found only on the UDSmr list of diagnoses. This "FIM only" variable thus measures whether the cases with the diagnosis found only on the FIM cost the same as the cases with the diagnosis found on the MEDPAR. These analyses included 22 of the 23 conditions in our final recommendation. Of the 44 "FIM only" coefficients for the 22 variables in our final recommendation, only two were significantly different from 0 at $p < 0.01$.

[2] Of course, HealthSouth's own database also includes the diagnoses found on the Uniform Bill discharge abstract. We received these diagnoses via the MEDPAR file and used them in all analyses.

comorbidities in our HealthSouth and COS data than in our UDSmr data. To reduce the effect of this bias on our coefficient estimates, in all the regressions reported here, we added dummy variables to indicate whether the source of the data was HealthSouth or COS. Consistent with our hypothesis of more frequent missing comorbidities in the HealthSouth and COS data, the source coefficient was always positive, statistically significant, and extremely small.

Constructing Hypotheses

A variety of hypotheses were generated by expert consultants to the project (Drs. Stineman and Zingmond), from the discussion in our 1997 study (originally from Drs. Rosenfeld and Rubenstein), and by a subcommittee of our TEP (Drs. Goldfield, Hoppe, Roth, Smith, and Stineman, and Mr. Hutchins and Ms. Davis). The form of each hypothesis was that a given condition would lead to an increase in the costs of rehabilitation, even after accounting for the patient's impairment and functional status. Many of the clinicians suggested the same or similar conditions as the cause of extra costs.

Some of the hypotheses were stated in terms of specific ICD-9-CM codes, others in terms of a clinical condition or group of similar conditions. The latter hypotheses were translated into ICD-9-CM codes using the hierarchical nature of the ICD-9-CM coding system.

An additional hypothesis was that the conditions that increase cost substantially in acute care would also increase cost in rehabilitation. In order to test this hypothesis, we turned to a preexisting list of codes, as we describe next.

Major Comorbidity Variable

The major comorbidities were first defined as those diagnoses found in a study of refined DRGs to have major effects on the costs of acute care stays. In our 1997 study, using 1994 discharges, we determined that these comorbidities also increased the cost of rehabilitation. We found that the effect of comorbidities varied across RICs, significantly increasing costs for patients in some RICs while having no effect on others. We further found that, in RICs where comorbidities affect cost, the amount of the cost increase in a RIC was as good a model as the percentage increase in the cost of each FRG within the RIC.

This major comorbidity list was coded using the ICD-9-CM diagnoses in use in FY 1994 (beginning October 1, 1993). We added new codes that are now used to

code patients who would have received a diagnosis based on the old major comorbidity list. We examined all the new codes published in the Federal Register for September 1994, 1995, 1996, and 1997 and for July 1998.[3] For each code that was a complication or comorbidity (CC) in one of the tables, we determined whether it was a further specification of a code that was a major comorbidity in the initial list and therefore would have been coded as a major comorbidity prior to creation of the new code. We created a dummy variable--1 if the patient had any secondary diagnosis on the major comorbidity list and 0 otherwise. In preliminary analyses, we verified that this set of diagnoses increased costs in CY 1996 and 1997, as it had done in 1994.

Defining Conditions

Using advice from our clinical consultants and the structure of the ICD-9-CM codes, we defined many candidate conditions. Each condition consisted of a list of ICD-9-CM codes and was suggested to us as specifically important in rehabilitation. Examples include dysphagia (difficulty in swallowing) and the complications of diabetes.

We partitioned the codes in the "major comorbidity" variable in order to take out extremely expensive conditions and other groups that had clinical coherence. In particular, we created separate groups for

- ventilator dependence
- meningitis and encephalitis
- selected candidiasis codes
- pneumonia
- infections.

Where necessary, we combined some new codes with the codes from the major comorbidity variable. For example, we added the code for candidal esophagitis to the candidiasis codes found in the "Major comorbidity" list.

Table 4.10, found at the end of this chapter, lists the ICD-9-CM codes that define each condition in our final recommendation.

In preliminary analyses, we created separate conditions for each of the types of diabetes complications: renal, ophthalmologic, circulatory, nervous system, and other. However, we found that the effects of all the complications

---

[3] No new codes were created for FY 2000 (FY 2000 includes the last quarter of CY 1999, which is in our data).

except renal were very similar--in the range of 0.07 to 0.09 more than for patients without any condition. Moreover, the variables were correlated: A reasonably large number of people had more than one type of complication, and cost was not related to whether more than one type of complication had been coded. This meant that, in a simultaneous regression model, individual coefficients on each type of disease were often not significant, though a single variable for the presence of any non-renal diabetes complication was significant. Consequently, in the analyses presented here we have only two variables for the complications of diabetes: renal manifestations and all other complications.

Similarly, in preliminary analyses, we used two variables for osteomyelitis: chronic osteomyelitis, and acute and not otherwise specified (NOS). However, chronic osteomyelitis is a relatively rare code and was not statistically significant by itself, although indistinguishable in effect from that of acute and NOS osteomyelitis. Consequently, we report on only a single osteomyelitis variable.

We analyzed two versions of the chronic skin ulcer condition. One is defined using diagnoses from the rehabilitation stay, just as for other conditions in our study. However, as a surrogate for chronic skin ulcer cases where the condition was present at admission, we defined another variable using the record for the acute care hospitalization that immediately preceded entry into the rehabilitation hospital. This "acute ulcer" condition includes only cases where skin ulcer was listed as a diagnosis during the acute care stay.

RIC Exclusions

One of the shortcomings of Carter, Relles, and Buchanan (1997) was that exactly the same set of diagnoses was evaluated in each RIC. Clearly, however, a comorbidity in one RIC might be a typical condition of the patients in another RIC. For example, stroke is clearly a complication in most RICs but will be present for all patients in RIC 1 whether coded or not. Similarly, cardiac conditions might complicate management in all RICs but should not be viewed as a comorbidity in the cardiac RIC 14.

We asked Drs. Stineman and Feinberg to review the list of ICD-9-CM codes in all conditions in our hypotheses and determine which diagnoses should not be counted as a comorbidity or complication in each RIC. Many diagnoses, including most infections, were not excluded from any RIC. A few diagnoses were excluded

from more than one RIC (e.g., meningitis codes were excluded from both RICs 3 and 5; brain hemorrhages were excluded from RICs 1, 2, and 3). The remaining diagnoses were excluded from only one RIC. (The exclusions are also listed in Table 4.10 and included as part of the definition of each condition.)

The RIC exclusions are numerically important and can strongly influence the measurement of the effect of conditions on cost. For example, 37 percent of stroke cases had at least one of the codes in the original major comorbidity list; after RIC exclusions were enforced, only 5.5 percent of cases had a major comorbidity code. The implementation of comorbidity exclusions substantially improved the ability of the remaining major comorbidities to predict cost in the stroke RIC.

Regressions

We used multiple regression to determine the effect of the hypothesized conditions on cost while controlling for FRG. Separate regressions were run for each year and for two pairs of years: the combined 1996-1997 and 1998-1999 data. The dependent variable for all regressions was the log of the wage-adjusted case cost. Each regression contained dummy variables for each FRG.

We first ran separate regressions for each RIC and for each hypothesized condition. This method was based on preliminary analyses where we found substantial, significant interactions between RIC and comorbidities.

As the RIC exclusions were improved and the comorbidity hypotheses grew to include several very rare conditions, we decided to base the statistical criteria for membership in the set of relevant conditions on a regression covering all RICs and using data from both 1998 and 1999. The regression included all conditions that were found to be positively related to cost in the preliminary regressions. Conditions were dropped if they were not significant at the 0.01 level or if their coefficient was 0.04 or lower. (Conditions that appeared marginal on these criteria were entered one at a time or in small groups into the regression containing the remaining conditions.) These criteria allow us to include rare conditions that have a substantial average effect on cost.

Certain conditions that were judged to be preventable in some cases were dropped from the final model, as discussed in more detail below.

Given the list of relevant conditions, we used regression to explore alternative payment structures, including the form of payment (a single payment,

a tiered structure like the refined DRGs, or an index) and whether the payment amount should vary across RICs.

CONDITIONS THAT DO AND DO NOT INCREASE COST

Conditions Recommended for Additional Payment

Table 4.2 shows the coefficient and t-statistic on each condition for four annual regressions of the logarithm of cost per case on FRG indicators and the listed conditions. The listed conditions consist of all those for which CMS specified an additional payment in the final regulation. The last column shows the pooled 1998-1999 regression. The criteria used to select this list of conditions were applied to regressions similar to this model but included additional conditions: significance at the 0.001 level or lower and a coefficient of 0.04 or higher. All the coefficients in the combined 1998-1999 regression are statistically significantly different from 0 at the 0.0001 level. In addition to significance and magnitude, another criterion for creating the list of conditions to be paid was policy acceptability. We discuss this criterion further below.

Table 4.2

Marginal Effect of Conditions in Final Model, Regression of Logarithm of
Wage-Adjusted Cost, by Year

Condition	1996 data		1997 data		1998 data		1999 data		1998-99 data	
	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.
Ventilator	0.7860	7.23	0.4182	5.43	0.2926	4.16	0.2331	4.52	0.2516	6.07
Misc. throat problems	0.1270[a]	1.26	0.1647	2.03	0.2300	2.95	0.2512	2.39	0.2371	3.78
Candidiasis (selected)	0.2019	3.91	0.1502	3.36	0.2144	4.90	0.2313	4.89	0.2219	6.91
Tracheostomy	0.2466	7.95	0.2029	7.31	0.2150	8.97	0.1735	7.21	0.1935	11.40
Vocal cord paralysis	0.2079	5.79	0.2169	6.25	0.2056	5.86	0.1607	5.03	0.1808	7.66
Malnutrition	0.0874	2.46	0.1336	4.25	0.1871	5.66	0.1467	4.53	0.1660	7.18
Intestinal infection clostridium	0.1439	7.03	0.1480	8.72	0.1475	9.39	0.1513	10.96	0.1497	14.43
Dialysis	0.1407	6.74	0.1986	10.93	0.1196	7.43	0.1574	10.59	0.1400	12.82
Pseudomonas	0.0506	2.97	0.1327	8.51	0.1170	7.30	0.1411	8.96	0.1294	11.52
Other infections	0.0929	9.57	0.1295	15.52	0.1239	15.82	0.1211	16.54	0.1225	22.90
Cachexia	0.0320[a]	0.64	0.1242	3.25	0.1703	4.52	0.0749	1.98	0.1228	4.60
Dysphagia	0.1363	11.26	0.1234	12.03	0.1187	12.21	0.1236	13.71	0.1214	18.37
Gangrene	0.1264	3.19	0.1268	3.80	0.1466	4.12	0.0814	2.49	0.1109	4.60
Meningitis and encephalitis	0.0885	2.40	0.0964	2.94	0.1544	4.79	0.0690	2.33	0.1069	4.91
Renal complications of diabetes	0.1212	8.08	0.1069	8.24	0.0992	8.00	0.1032	8.75	0.1010	11.81
Hemiplegia	0.0657	7.89	0.0532	6.58	0.0891	7.69	0.0977	8.24	0.0933	11.27
Aplastic anemia and selected anemias	0.0825	2.87	0.1096	4.50	0.0800	3.27	0.0979	4.29	0.0896	5.37
"Major" comorbidities	0.0755	11.98	0.0826	14.14	0.0891	15.91	0.0873	16.50	0.0881	22.91
Obesity	0.0596	2.90	0.0445	2.63	0.0991	6.53	0.0628	4.47	0.0794	7.70
Esophagal conditions	0.1133	3.88	0.0309[a]	1.11	0.0574	2.24	0.0786	3.33	0.0687	3.96
Pneumonia	0.0885	8.22	0.0639	6.72	0.0673	7.73	0.0556	7.17	0.0607	10.48
Non-renal complications of diabetes	0.0323	4.34	0.0493	7.42	0.0570	8.73	0.0526	8.47	0.0545	12.12
Amputation of LE	0.0335[a]	1.70	0.0502	3.08	0.0536	3.41	0.0368	2.47	0.0453	4.19
R-squared	0.3586		0.3551		0.3424		0.3358		0.3386	

[a] $p > 0.05$.

Note: Each year's regression controlled for each FRG; the 1998-1999 also controlled for year. All coefficients are significant with $p < 0.05$, unless otherwise noted.

The coefficients can be loosely interpreted as the fraction by which the cost of a case with the condition exceeds the cost of a case in the same FRG without this condition, but with exactly the same list of other conditions in the model.[4]

All the conditions have a substantial effect on the cost of a case. Using the combined 1998-1999 data, for instance, we see that ventilator status increases costs by over 25 percent. The infrequent "miscellaneous throat problems" diagnosis (usually a foreign body in the larynx) increases cost by about 24 percent. The smallest effect, at 4.5 percent, is found with amputation of the lower extremity--which, of course, is excluded from the RIC of the same name.

There is substantial year-to-year variation in many of the coefficients. Much of the variation is due to random fluctuation because the sample sizes for many of these conditions are very small. Table 4.3 shows the percentage of cases coded with each condition in each year. For example, in 1996 only 0.015 percent, or 15 cases out of 100,000, were coded as ventilator dependent. In addition to random fluctuation, some of the variance may be due to changes in coding practices. The average absolute value of the difference between the 1996 and 1997 coefficients is 0.0464.[5] The difference between the 1997 and 1998 coefficients declined by one-third, to 0.0320. The difference declined again by 10 percent in 1998-1999. Thus the measured effects are becoming more stable over time, and we think some of the increased stability is due to improved coding. Further, there may be real trends in the relative costs of care for patients with some conditions. For example, it may be that as ventilator-dependent patients become more frequent in some hospitals, their care is becoming more standardized and more efficient. As coding improves further, we would expect the difference between coefficients on annual regressions to asymptotically reach a value that is due solely to random fluctuation and trends, if any.

---

[4] A more precise statement is that the ratio of the expected cost of a case with the condition to the expected cost of an otherwise similar case without the condition is the exponential of the coefficient. For small values of x, the exponential of x is close to 1 + x. For example, if the coefficient of a condition is x = 0.10, the ratio of costs is 1.1052, so the fraction by which the cost for a case with the condition exceeds the cost for an otherwise similar case without the condition is 0.1052.

[5] Each coefficient is counted equally in this average.

Table 4.3

**Percentage of All Cases That Have Each Condition in the
Final Model, by Year**

Condition	Percentage of Cases with Condition			
	1996	1997	1998	1999
Ventilator	0.015	0.025	0.027	0.047
Miscellaneous throat problems	0.017	0.022	0.022	0.011
Candidiasis (selected)	0.065	0.073	0.070	0.055
Tracheostomy	0.183	0.192	0.238	0.218
Vocal cord paralysis	0.135	0.121	0.110	0.121
Malnutrition	0.136	0.148	0.123	0.117
Intestinal infection clostridium	0.414	0.511	0.550	0.652
Dialysis	0.402	0.449	0.527	0.567
Pseudomonas	0.600	0.603	0.526	0.498
Other infections	2.005	2.173	2.273	2.382
Cachexia	0.069	0.100	0.095	0.086
Dysphagia	1.254	1.484	1.516	1.607
Gangrene	0.110	0.132	0.106	0.115
Meningitis and encephalitis	0.127	0.136	0.130	0.141
Renal complications of diabetes	3.706	3.947	3.739	3.786
Hemiplegia	2.748	2.436	1.086	0.948
Aplastic anemia and selected anemias	0.210	0.247	0.225	0.236
"Major" comorbidities	5.049	4.689	4.690	4.812
Obesity	0.410	0.515	0.587	0.626
Esophageal conditions	0.203	0.190	0.206	0.221
Pneumonia	1.525	1.663	1.832	2.123
Non-renal complications of diabetes	0.941	1.060	1.070	1.073
Amputation of lower extremity	0.462	0.572	0.564	0.577
Number of cases	128,591	153,317	172,136	189,523

Other Conditions That Increase Cost

Some conditions that meet statistical criteria for increasing cost were not included as warranting extra payment. They fall into a group we call "possibly preventable conditions" to indicate that some clinicians believe they are preventable, at least in some cases. These include urinary tract infections, chronic skin ulcers, selected thrombophlebitis codes, and osteomyelitis. Table 4.4 shows regressions similar to those in Table 4.2, using 1998 and 1999 data[6] but adding these four conditions. All are highly significant in each year. In the combined 1998-1999 data, cases with one of these conditions cost from 10 percent to 16 percent (= 100 *(exp (0.1482)-1)) more than otherwise similar cases without the condition.

---

[6] For the sake of brevity, we omit the 1996 and 1997 regressions. However, all four variables were significant in each year (p < 0.01) and had magnitudes roughly similar to those in the 1998-1999 data.

Table 4.4

**Marginal Effect of All Conditions Found to Increase Cost,
Regression of Logarithm of Wage-Adjusted Cost, by Year**

Condition	1998 data		1999 data		1998-99 data	
	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.
Thrombophlebitis	0.1322	13.81	0.1619	18.16	0.1482	22.71
Chronic skin ulcer	0.1292	19.70	0.1335	21.46	0.1314	29.12
Osteomyelitis	0.1151	5.60	0.1110	5.91	0.1127	8.14
Urinary tract infection	0.0953	26.59	0.1000	28.52	0.0977	38.97
Ventilator	0.3146	4.49	0.2348	4.58	0.2598	6.29
Miscellaneous throat problems	0.2366	3.05	0.2569	2.45	0.2454	3.93
Candidiasis (selected)	0.2064	4.73	0.2171	4.61	0.2113	6.61
Tracheostomy	0.2159	9.04	0.1722	7.18	0.1936	11.45
Vocal cord paralysis	0.2070	5.93	0.1636	5.14	0.1826	7.76
Malnutrition	0.1816	5.52	0.1252	3.88	0.1528	6.63
Intestinal infection clostridium	0.1323	8.44	0.1324	9.63	0.1323	12.80
Dialysis	0.1250	7.79	0.1594	10.77	0.1436	13.20
Pseudomonas	0.0490	3.04	0.0703	4.43	0.0599	5.30
Other infections	0.1150	14.72	0.1113	15.24	0.1130	21.19
Cachexia	0.1645	4.38	0.0680	1.81	0.1156	4.35
Dysphagia	0.1192	12.30	0.1221	13.60	0.1209	18.36
Gangrene	0.1206	3.40	0.0471	1.44	0.0804	3.35
Meningitis and encephalitis	0.1437	4.47	0.0634	2.15	0.0994	4.58
Renal complications of diabetes	0.1022	8.27	0.1043	8.88	0.1033	12.13
Hemiplegia	0.0909	7.87	0.0993	8.42	0.0951	11.52
Aplastic anemia and selected anemias	0.0746	3.06	0.0917	4.03	0.0839	5.05
"Major" comorbidities	0.0833	14.92	0.0815	15.46	0.0824	21.49
Obesity	0.0983	6.50	0.0579	4.14	0.0765	7.44
Esophagal conditions	0.0579	2.27	0.0830	3.53	0.0712	4.12
Pneumonia	0.0658	7.59	0.0546	7.08	0.0594	10.31
Non-renal complications of diabetes	0.0494	7.57	0.0436	7.03	0.0462	10.29
Amputation of LE	0.0524	3.35	0.0360	2.43	0.0440	4.09
R-squared	0.3475		0.3417		0.3444	
Skin ulcer in acute care stay	0.0598	7.08	0.0920	12.07	0.0776	13.72
R-squared	0.3426		0.3363		0.3392	

Note: Each year's regression controlled for each FRG. All coefficients are significant with $p < 0.05$.

There was substantial disagreement among members of the TEP subcommittee on comorbidities about whether facilities should be paid an extra amount for cases with possibly preventable complications. There was even disagreement about whether one could be sure that these conditions were sometimes actually preventable. For possibly preventable conditions, there is a tension between being fair to hospitals that get more patients who are prone to these conditions and rewarding hospitals that provide poor care. Because of budget neutrality, an explicit payment for the conditions would penalize hospitals that invest in effective preventive activities at a greater-than-average rate.

The possibly preventable conditions are strongly concentrated in specific FRGs. Table 4.5 shows the percentage of cases in each RIC with each condition. It also shows the percentage of cases with the condition in the lowest-weighted FRG, which is the FRG with the highest-functioning patients. Similarly, the percentage of cases in the highest-weighted FRG gives the percentage of cases with the condition among patients with the least function.

Urinary tract infection (UTI) is highly concentrated in specific RICs and in persons with low function at admission. As shown in Table 4.5, UTI is relatively prevalent in the inpatient rehabilitation population; it affects 13.5 percent of all cases.[7] The UTI cases are found more frequently in brain and spinal cord injury patients and those with stroke or Guillain-Barré and less frequently in the amputation, pulmonary, cardiac, and major joint replacement RICs. However, function is at least equally important in identifying patients prone to UTI. As one can see from Table 4.5, within each RIC, a much higher proportion of patients have UTI in the highest-weight (lowest-functional-independence) FRG than in the FRG with the lowest weight. Overall, only 6.4 percent of cases in the FRG with the lowest weight in each RIC have UTI whereas 20.4 percent of cases in the highest-weighted FRG in each RIC have UTI. Looking across RICs, the rate of UTI varies from a low of 1.9 percent (for the most-independent group of patients in the pulmonary RIC) to 35.8 percent (for the least-independent patients with traumatic spinal cord injury).

---

[7] The numbers of cases by RIC differ from those in Table 4.1 in that Table 4.5 includes transfer cases. Payment for transfer cases will be affected by decisions concerning comorbidities--although we do not use transfer cases to define the FRGCs.

Table 4.5

Percentage of 1998 and 1999 Cases with Chosen Conditions for Each RIC and for
Low-Weighted and High-Weighted FRGs

RIC	Number of cases	Urinary Tract Infection (%)			Chronic Ulcers (%)			Thrombophlebitis (selected codes) (%)		
		Complete RIC	Lowest-weight FRG	Highest-weight FRG	Complete RIC	Lowest-weight FRG	Highest-weight FRG	Complete RIC	Lowest-weight FRG	Highest-weight FRG
Stroke	108,090	16.9	4.8	25.5	2.4	0.6	4.8	1.7	0.3	3.2
Traumatic BI	5,989	16.5	7.2	21.4	2.9	1.3	4.9	3.2	1.0	5.0
Nontraumatic BI	10,315	17.9	9.7	26.0	3.3	0.8	6.7	3.9	1.8	6.9
TSCI	2,818	26.5	14.4	35.8	15.3	8.8	21.3	2.5	1.3	3.9
Nontraumatic SCI	14,404	19.3	9.4	31.1	5.7	1.8	13.4	2.3	0.9	4.1
Neurological	22,309	15.2	7.3	22.4	5.5	2.0	9.1	1.6	0.6	2.3
Hip fracture	54,447	16.3	9.5	20.8	4.6	1.9	7.3	1.7	1.4	2.1
Major joint replacement, LE	93,569	9.5	5.0	17.3	1.7	0.5	6.2	1.5	1.2	2.0
Other orthopedic	22,660	12.4	8.3	18.4	2.8	1.1	5.3	1.2	0.7	1.6
Amputation, LE	16,324	9.1	3.3	14.3	14.2	8.1	19.9	1.1	0.8	1.5
Amputation, other	1,726	5.6	2.7	8.0	15.1	13.3	18.6	1.0	0.6	2.1
Osteoarthritis	10,926	10.1	6.2	15.1	2.7	0.9	5.9	1.4	1.2	1.5
Other arthritis	5,446	10.3	6.1	15.1	4.3	1.6	9.3	1.4	1.1	2.0
Cardiac	19,136	9.3	5.5	15.0	4.1	1.6	8.7	1.6	1.2	2.2
Pulmonary	12,757	6.7	1.9	12.9	3.8	1.3	8.7	1.5	0.7	2.3
Pain syndrome	6,847	11.3	8.4	16.8	2.0	1.1	3.7	1.0	0.7	1.7
MMT, no BI, SCI	4,407	15.4	10.2	21.9	3.6	1.7	6.2	2.0	1.5	3.0
MMT, with BI or SCI	739	19.1	8.6	25.8	5.0	1.4	6.8	4.7	0.0	5.9
Guillain-Barré	809	20.3	10.0	32.4	4.5	2.0	8.1	3.8	1.2	5.3
Miscellaneous	52,965	13.4	7.9	18.2	7.4	4.2	13.2	2.7	1.8	3.8
Burns	309	15.2	8.0	19.4	6.1	6.2	6.1	5.5	8.8	3.6
All	466,992	13.5	6.4	20.4	4.1	1.7	7.9	1.8	1.0	2.7
Range across FRGs			1.9	35.8		0.5	21.3		0.0	6.9

Table 4.5 (cont.)

RIC	Number of cases	Osteomyelitis (%)			Chronic Ulcer in Acute Stay (%)			Any Complication (%)		
		Complete RIC	Lowest-weight FRG	Highest-weight FRG	Complete RIC	Lowest-weight FRG	Highest-weight FRG	Complete RIC	Lowest-weight FRG	Highest-weight FRG
Stroke	108,090	0.1	0.0	0.1	1.2	0.6	2.4	19.9	5.6	30.8
TBI	5,989	0.1	0.0	0.1	1.4	0.7	2.5	21.0	9.5	28.4
Nontraumatic BI	10,315	0.1	0.1	0.2	1.7	0.8	3.0	23.1	12.1	34.9
TSCI	2,818	1.8	1.3	1.4	9.9	7.2	13.4	38.0	20.8	50.3
Nontraumatic SCI	14,404	1.1	0.3	2.1	3.1	1.2	7.0	25.2	11.8	42.7
Neurological	22,309	0.4	0.4	0.4	4.0	2.1	6.2	20.8	10.0	30.7
Hip fracture	54,447	0.1	0.2	0.1	1.6	1.0	2.5	21.0	12.4	27.5
Major joint replacement, LE	93,569	0.2	0.1	0.2	0.6	0.2	2.0	12.3	6.7	23.6
Other orthopedic	22,660	0.7	0.5	1.0	1.7	0.9	2.9	16.1	10.2	24.2
Amputation, LE	16,324	2.5	2.6	2.2	13.0	6.4	15.6	24.2	13.6	33.3
Amputation, other	1,726	6.6	5.8	5.4	21.1	20.3	20.0	25.3	21.2	28.8
Osteoarthritis	10,926	0.3	0.4	0.6	1.5	0.5	3.6	13.5	8.4	21.0
Other arthritis	5,446	0.8	0.7	1.2	3.0	1.2	6.8	15.6	9.4	24.6
Cardiac	19,136	0.2	0.1	0.3	2.5	1.1	5.5	14.1	8.2	24.1
Pulmonary	12,757	0.1	0.0	0.1	2.6	0.6	5.6	11.2	4.0	21.1
Pain syndrome	6,847	0.5	0.3	1.0	1.1	0.7	1.7	14.2	10.1	21.8
MMT, no BI, SCI	4,407	0.2	0.1	0.2	1.6	1.0	1.9	19.9	12.8	28.8
MMT, with BI or SCI	739	0.7	0.0	0.0	2.7	1.4	4.5	25.8	10.0	33.9
Guillain-Barré	809	0.1	0.0	0.0	2.5	0.4	4.5	26.1	12.5	40.5
Miscellaneous	52,965	0.6	0.5	0.7	5.5	3.3	9.8	21.8	13.6	31.1
Burns	309	0.0	0.0	0.0	5.5	6.2	5.1	23.6	20.4	25.5
All	466,992	0.4	0.3	0.5	2.5	1.3	4.4	18.4	9.1	28.3
Range across FRGs			0.0	5.4		0.2	20.0		4.0	50.3

Note: Based on 1998 and 1999 data. Excludes cases with LOS less than or equal to three days and in-hospital deaths.

Similarly, chronic skin ulcer cases are concentrated by RIC and function. Patients with traumatic spinal cord injury and amputations are much more likely to suffer from decubitus ulcers or chronic leg ulcers. Overall, only 1.7 percent of cases in the FRG with the lowest weight in each RIC have such ulcers, while 7.9 percent of cases in the highest-weighted FRG have ulcers. In all RICs except burns, function is highly related to the presence of ulcers. Looking across RICs, the rate of ulcers varies from 0.5 percent (for the most-independent major joint replacement patients) to 21.3 percent (for the least-independent TSCI patients).

Cases with selected thrombophlebitis codes are also concentrated, both by RIC and by function. However, low incidence means that there are no FRGs that have more than 6.9 percent of cases with this condition. Osteomyelitis cases are concentrated in the amputation RICs, where there appears to be no relationship to function. The low incidence of osteomyelitis in other RICs masks the pattern of incidence of this disease, if any, in our data.

The high concentration of these diseases in specific FRGs means that the payment for a typical case in such an FRG will be noticeably higher if we do NOT pay for these diseases than if we do. This would reward hospitals that invest in preventive activities. In addition to increasing incentives, this payment would provide the funds for preventive activities. Nevertheless, several of our TEP subcommittee members believed that these cases should be paid for in proportion to their extra cost. They had concerns about financial risk to hospitals, about how frequently these conditions can really be prevented, and about incentives to diagnose and treat the conditions.

The disagreement in the TEP concerned only conditions that arose (or were identified) following the assessment period. There was unanimity in the subcommittee that CMS should pay for conditions present at admission to the rehabilitation hospital if they require treatment or affect treatment in such a way as to increase cost. Our data on comorbidities and complications in the rehabilitation stay do not tell us when the condition arose. In order to determine if chronic ulcers might be important when present at admission, we defined a variable telling whether a chronic skin

ulcer was noted during the acute care hospitalization that preceded the rehabilitation stay.

As seen in Table 4.5, only 2.5 percent of cases had such a code, compared to the 4.1 percent of cases for whom an ulcer was noted during the rehabilitation stay. We added this variable alone to the regression in Table 4.2 and found that inpatient rehabilitation costs for cases with chronic skin ulcers noted during the acute stay were about 8 percent higher than otherwise similar cases without such a notation (or without an immediately preceding acute stay). This is about 60 percent of the extra cost associated with ulcers noted during the rehabilitation stay. It is not clear whether the difference is due to treatment received in acute care or inaccuracy of coding in either the acute stay or the rehabilitation stay.

We have no record of whether the IRF found the possibly preventable complications at admission. Consequently, we have no way of estimating the extra cost of such conditions. Because we have no estimate of incidence, we have no way of estimating the effect of paying for such conditions on budget neutrality. Therefore, we cannot recommend any such payment in the initial year of the IRF PPS. However, the new IRF patient assessment instrument (PAI) does request information about diagnoses arising during the hospitalization. Although the form of the request is unfortunate because it requires double entry of some diagnoses, if the rules are carefully followed we may be able to distinguish diagnoses present at admission from those that arise during the IRF stay. If cases presenting at admission with these conditions do cost more, we will be able to consider such a payment in the first refinement phase.

We found that skin ulcers present at admission are likely to be widespread (probably more than 2.5 percent of all cases) and may have an important effect on the costs of a case (between 8 and 14 percent, depending on why our proxy measure for ulcers present at admission resulted in lower costs than ulcers noted during the rehab stay). Therefore, the issue of complications present at admission is important, and we believe that this issue should have high priority in initial efforts to refine the comorbidity payment.

Cancers Do Not Cost More

Certain hypotheses were repeated continually by all of our
consultants. In this section we investigate one of these hypotheses: that
cases with specific cancers should cost more than cases without cancer,
after controlling for function and other conditions known to increase
cost. The specific cancers that were tested were lung cancer, brain
cancer, cancer of the blood or blood forming organs, and metastasized
cancer.

Table 4.6 shows the condition coefficients in annual regressions of
the log of costs on FRG, the four conditions being tested, and all the
conditions in our final model. We also added a variable for "other
cancers" from which we removed all codes for cancers in remission; the
codes for cancers in remission were entered as a separate variable for
completeness. The clear conclusion is that patients with these cancers do
not cost more than otherwise similar patients without cancer. Exactly half
of the 16 annual coefficients for the four hypotheses were negative. Only
two coefficients were significantly different than 0--one positive and one
negative.

The positive and significant coefficient is the 1997 coefficient for
cancers of the blood. In preliminary analyses, based only on the 1996 and
1997 data, we examined subsets of this infrequent disease category and
reached the conclusion that patients with Hodgkin's disease cost more than
other patients without a relevant comorbidity. However, in the 1998 and
1999 data, patients with Hodgkin's disease did not cost more than
otherwise similar patients without a relevant comorbidity--and neither did
all patients with a blood-related cancer. We believe that our finding in
the 1997 data about Hodgkin's disease was a statistical anomaly that arose
because we had divided the data into groups that were too small.

Table 4.6

Marginal Effect of Cancers and Variables in Final Model,
Regression of Logarithm of Wage-Adjusted Cost, by Year

Condition	1996 Data		1997 Data		1998 Data		1999 Data		1998-1999 Data	
	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.
Lung cancer	-0.0575	-2.81	-0.0285[a]	-1.52	0.0189[a]	1.07	-0.0174[a]	-1.08	-0.0012[a]	-0.10
Brain cancer	0.0310[a]	0.50	-0.0638[a]	-1.10	0.0086[a]	0.17	-0.0342[a]	-0.57	-0.0091[a]	-0.24
Metastasized cancer	-0.0016[a]	-0.10	-0.0181[a]	-1.31	-0.0041[a]	-0.30	0.0015[a]	0.11	-0.0012[a]	-0.13
Cancer of blood or blood organs	0.0043[a]	0.26	0.0390	2.60	0.0139[a]	0.98	0.0128[a]	0.94	0.0134[a]	1.36
Cancer in remission	0.1214[a]	1.91	0.0361[a]	0.57	0.0258[a]	0.42	0.0438[a]	0.79	0.0361[a]	0.87
Other cancer	-0.0100[a]	-0.91	-0.0069[a]	-0.67	-0.0199	-2.10	-0.0230	-2.56	-0.0216	-3.31
Ventilator	0.7884	7.25	0.4182	5.43	0.2923	4.16	0.2334	4.53	0.2510	6.05
Miscellaneous throat problems	0.1264[a]	1.26	0.1655	2.04	0.2292	2.94	0.2518	2.39	0.2386	3.81
Candidiasis (selected)	0.2023	3.91	0.1496	3.35	0.2145	4.90	0.2319	4.90	0.2223	6.92
Tracheostomy	0.2472	7.97	0.2025	7.29	0.2159	9.00	0.1736	7.21	0.1943	11.44
Vocal cord paralysis	0.2094	5.83	0.2165	6.23	0.2057	5.87	0.1616	5.06	0.1808	7.66
Malnutrition	0.0877	2.46	0.1333	4.24	0.1874	5.67	0.1481	4.57	0.1671	7.22
Intestinal infection clostridium	0.1439	7.03	0.1474	8.69	0.1476	9.39	0.1514	10.96	0.1496	14.43
Dialysis	0.1405	6.73	0.1983	10.91	0.1194	7.41	0.1569	10.55	0.1396	12.78
Pseudomonas	0.0504	2.96	0.1328	8.51	0.1168	7.29	0.1411	8.96	0.1292	11.50
Other infections	0.0931	9.59	0.1294	15.50	0.1239	15.82	0.1210	16.53	0.1224	22.89
Cachexia	0.0387[a]	0.77	0.1251	3.28	0.1701	4.51	0.0754	1.99	0.1221	4.58
Dysphagia	0.1360	11.23	0.1234	12.04	0.1188	12.21	0.1236	13.70	0.1215	18.37
Gangrene	0.1264	3.19	0.1267	3.80	0.1461	4.10	0.0809	2.47	0.1103	4.58
Meningitis and encephalitis	0.0882	2.39	0.0954	2.91	0.1540	4.77	0.0687	2.32	0.1071	4.92
Renal complications of diabetes	0.1210	8.06	0.1067	8.22	0.0991	7.98	0.1028	8.72	0.1009	11.81
Hemiplegia	0.0656	7.88	0.0534	6.61	0.0889	7.65	0.0983	8.29	0.0935	11.27

Table 4.6 (cont.)

Condition	1996 Data		1997 Data		1998 Data		1999 Data		1998-1999 Data	
	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.
Aplastic anemia and selected anemias	0.0835	2.90	0.1090	4.47	0.0791	3.23	0.0985	4.30	0.0895	5.35
"Major" comorbidities	0.0754	11.96	0.0826	14.12	0.0891	15.90	0.0872	16.49	0.0881	22.90
Obesity	0.0594	2.89	0.0444	2.63	0.0990	6.52	0.0624	4.44	0.0792	7.68
Esophagal conditions	0.1130	3.87	0.0316[a]	1.14	0.0574	2.24	0.0786	3.33	0.0686	3.95
Pneumonia	0.0885	8.21	0.0637	6.70	0.0672	7.72	0.0555	7.15	0.0605	10.45
Non-renal complications of diabetes	0.0321	4.30	0.0491	7.39	0.0570	8.72	0.0525	8.45	0.0545	12.11
Amputation of LE	0.0334[a]	1.70	0.0500	3.07	0.0536	3.42	0.0364	2.45	0.0448	4.15
R-squared	0.3586		0.3552		0.3424		0.3358		0.3389	

[a] $p > 0.05$.

Note: Each regression controlled for each FRG and for year.

Although we do not show the data in detail, if we enter a single cancer variable instead of the six cancer variables shown in Table 4.6, we find that, in each of the four years of data, patients with cancer cost slightly less on average than otherwise similar patients without cancer. The coefficient is significant in the combined 1998-1999 data ($p = 0.04$). It is likely that some fraction of cancer patients are terminally ill. It makes sense for these patients to return to their home as quickly as possible so they can have time at home with their families. This reduction in cost appears to offset whatever increased costs, if any, are incurred for other patients.

PAYMENT

In this subsection we consider alternative ways to determine payment amounts for cases with the conditions listed in Table 4.2. An obvious alternative is to use the coefficients from a regression like the one found in Table 4.2 as if they were the marginal cost of each disease. A weight would then be calculated for each patient by multiplying the weight for a person in the patient's FRG with no comorbidities by the exponential of the coefficients for each condition that the patient has. So the weight for patients with two conditions would increase by two factors, the weight for those with three conditions by three factors, and so on.

Although a payment model based on the specifications in Table 4.2 might possibly be the most accurate available representation of the extra costs associated with the conditions, it has several drawbacks. First, because many of the conditions in Table 4.2 are very rare, the measured costs of those conditions are highly uncertain. Indeed, as discussed above, many coefficients change substantially from year to year. Yet we want a payment formula to be stable from year to year and not substantially affected by random events. Second, having additive coefficients for each condition may not lead to the best prediction of costs. As we show below, although there are substantial differences across RICs in the effect of many comorbidities and in the effect of groups of comorbidities, many of these differences cannot be reliably modeled at the condition level in our data. For many of the conditions listed in Table 4.2, there are not enough sample cases to estimate RIC-specific effects in

most RICs--predictions would have very high standard errors. Thus, the use of additive coefficients for each condition might force us either to ignore the interaction of RIC and comorbidity or to drop some conditions from some RICs. The result would not be necessarily the best prediction of costs.

Another important reason to be concerned with the model in Table 4.2 as a payment model relates to coding accuracy and budget neutrality. As discussed in the introduction, Congress mandated that payments under the rehabilitation PPS in FY 2002 should equal what payments would have been under the current payment system. It is very likely that comorbidities are undercoded in our data because these data fields were never used for payment. If payment were increased for each comorbidity that was coded, there would be a much greater risk of overpaying under the rehabilitation PPS than if each patient were limited to at most one comorbidity payment. Allowing each patient at most one comorbidity payment also eliminates the incentive to "game the system" by adding multiple diagnoses--although not the incentive to add a first relevant comorbidity that did not actually affect care.

Alternative Payment Strategies

We have found repeatedly (and demonstrate again here) that the extra cost of comorbidities varies across RICs. For example, most lower extremity joint replacement patients (RIC 8) are quite healthy (indeed, good health is often a prerequisite for this elective surgery). Consequently, their rehabilitation is accomplished quickly and LOS is typically much shorter than for other patients. However, patients in RIC 8 with a relevant comorbidity are indeed quite sick and so often require a much greater increase in resources (including LOS) than do patients in other RICs where the patients without a comorbidity have much greater needs. Because the extra cost of comorbidities varies across RICs, most of the payment strategies we examined allow payment increases for cases with comorbidities to differ across each RIC, according to how relative costs vary across RICs.

We considered only strategies where the payment amount for a case with a comorbidity is a multiple of the payment for a patient in the same

FRG with no relevant comorbidity. There are not enough data to establish separate comorbidity payment amounts for each FRG. The data that we do have are consistent with the fact that the increase in costs due to comorbidities is the same percentage across all FRGs in the same RIC.[8] Further multiplicative increases appear to make sense. First, most of the comorbidities require additional services to be delivered periodically (e.g., dialysis, respiratory services, drugs, laboratory tests), and thus per diem costs are a fixed percentage higher than those of other patients. Since much of the variation in costs across FRGs within the same RIC is a variation in LOS, the increased per diem costs are well approximated by a fixed-percentage increase in costs per case. Second, comorbidities can also slow the progress of rehabilitation. If the slowing is at a fixed rate per day within an RIC, we would again expect a fixed-percentage increase in cost.

The simplest possible comorbidity payment strategy would be to increase payment by a single fixed percentage for patients with any relevant comorbidity. The increased payment would be the estimated percentage by which the average cost of a case with at least one of the comorbidities exceeded the average cost of a case in the same FRG with none of the comorbidities. The NPRM suggested using a percentage increase for a set of relevant conditions, with the amount of the increase varying by RIC (see HCFA, 2000). We first evaluate the NPRM proposal.

As shown in Table 4.2, there are more conditions that affect cost than those that were found in the research leading up to the NPRM. Consequently, we also evaluate the strategy of the NPRM (a percentage increase in cost that varies only by RIC) after defining a relevant comorbidity as any of the conditions in Table 4.2 (more completely defined as any of the codes listed in Table 4.10).

Using a single comorbidity variable has the advantage of simplicity. In addition, it may increase stability over time, reduce the effect of coding improvements on budget neutrality, and reduce gaming. However, because it averages across conditions that have varying cost effects, such

---

[8] The interaction of the recommended comorbidity tier variable (with four levels) and FRG is significant at the $p = .01$ level only in RIC 20.

a system underpays cases with the most-expensive comorbidities and overpays those with the least-expensive comorbidities. Figure 4.1 plots the coefficients from the 1998-1999 regression in Table 4.2 and adds a two-standard-error confidence interval around each coefficient. Although, as discussed earlier, we cannot measure the effect of rare conditions very well, Figure 4.1 shows that we can be relatively certain that most of the conditions with larger coefficients (say, greater than 0.15) actually cost more than most of the conditions with small coefficients (say, 0.10 or less). This suggests that we may retain some of the advantages of the single comorbidity payment but still improve accuracy if we use a "tiered" approach that provides a small number of payment rates depending on the patient's comorbidities.

To evaluate the tiered approach, we divided the conditions in Table 4.2 into three groups based on their coefficients in the 1998-1999 regression: 0.16 or higher, 0.11 to 0.15, and 0.10 or lower (all rounded to two digits). To keep the advantages of a single comorbidity payment per patient, each patient would be assigned to a single tier. A patient with more than one comorbidity would be assigned to the tier containing the most expensive comorbidity. We recommended this tiered approach to CMS for use in the IRF PPS.

To understand the cost in accuracy of maintaining only one comorbidity payment per patient, we also considered more complicated alternative payment schemes, including a set of three indices, each of which counts the number of the patients' conditions in one of the tiers. Finally, we compared the tiered approach to using payment based on the multivariate model of Table 4.2, both with and without RIC-specific effects. Although we do not advocate using either of these models for the initial payment system because of incomplete coding in our data set, a more complicated payment scheme may be appropriate given the more accurate and complete data that will be available under the IRF PPS.
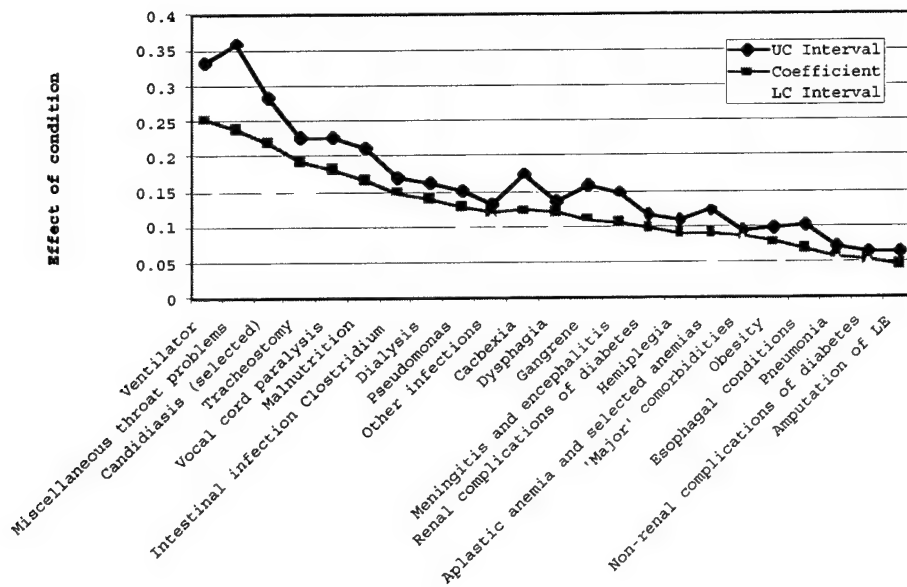
**Figure 4.1--Coefficient and Confidence Interval for
Effect of Condition on Log-Cost**

Evaluation of Payment Alternatives

Table 4.7 shows how accurately ten alternative regression models
predict the log of wage-adjusted cost per case. The first row of the table
(Model 1) gives the results with no comorbidity adjustment at all. Models
2 through 6 include different comorbidity variables, each one interacted
with each RIC. The F-statistics on the model line show that the
interaction of the set of comorbidity variables with RIC is highly
significant.

The R-squared column gives a measure of the goodness of fit of each
regression model. Cases with one of the relevant comorbidities are very
rare; consequently, each comorbidity variable adds relatively little to
the overall fit of the model. However, because cases with comorbidities at
admission are easily recognized and because the magnitude of the increase
in expected cost is quite large, it is important to pay adequately for

Table 4.7

Summary Statistics for Alternative Comorbidity Regression Models
(pooled 1998-1999 data)

Model and comorbidity variables	Avg. Value of Variable	R-squared	F for Interaction of Comorbidity Variable(s) and RIC			Percentage of Possible Comorbidity Effect Captured by Model
			Value	P	Degrees of Freedom	
1. No comorbidity effect	NA	0.3322	NA	NA	0	0.0
2. Single variable as defined in NPRM	0.0875	0.3352	13.92	0.0001	20	35.3
3. Single variable defined for Table 4.2 conditions	0.1680	0.3379	17.18	0.0001	20	68.9
4. Tier assignment		0.3388	6.57	0.0001	60	79.4
Tier 1	0.0056		2.87	0.0001	20	
Tier 2	0.0551		6.42	0.0001	20	
Tier 3	0.1072		11.26	0.0001	20	
5. Count conditions		0.3392	7.16	0.0001	60	84.9
Tier 1	0.0058		1.91	0.0084	20	
Tier 2	0.0589		5.56	0.0001	20	
Tier 3	0.1410		13.10	0.0001	20	
6. Individual conditions	NA	0.3405	2.53	0.0001	412	100.0
Positive & significant coef. only						
7. Single tier assignment	NA	0.3388	7.05	0.0001	50	85.8
8. Count conditions	NA	0.3392	8.06	0.0001	48	91.6
9. Individual conditions	NA	0.3398	4.41	0.0001	157	100.0
10. Individual conditions, no interaction with RIC	NA	0.3386	NA	NA	0	83.6

Note: Each regression controls for FRG and year in addition to shown comorbidities. NA = not applicable.

these cases to insure access. Further, because the admission policies of some rehabilitation hospitals may, for reasons other than PPS reimbursement, result in quite different distributions of these cases across hospitals, fairness also requires that CMS pay adequately for comorbidities. Although the absolute change in the R-squared may not measure all the value of comorbidity payments to the system, the R-squareds of alternative ways of paying for comorbidities do show their relative ability to match expected payments to expected costs.

The highest R-squared in the table is for the model that interacts each of the individual conditions in Table 4.2 with each RIC (Model 6). This model thus provides the smallest error in the prediction of costs that is possible with this definition of relevant comorbidity. The last column of the table expresses the increase in R-squared from Model 1 to the following models as a percentage of the increase from Model 1 to Model 6. The NPRM comorbidity model captured only 35.3 percent of the comorbidity effect that is possible with the set of comorbidities in Table 4.2. Using the NPRM strategy with the Table 4.2 set of comorbidities gets us 68.9 percent of the increase possible using this definition of relevant comorbidities. Using the tiered approach provides a further improvement in fit compared to a single comorbidity variable. Counting conditions within each tier gets us 84.9 percent of the way to having separate additive factors for each comorbidity.

The first six models in Table 4.7 include all variables regardless of whether they are significantly different from 0--or even positive. A payment system cannot include cost estimates that are just random effects. A simple way to avoid randomness in the payment system would be to just drop insignificant coefficients along with negative ones.[9] This is done for the tier model, the count conditions model, and the individual

---

[9] We will later propose a different method of dealing with marginally insignificant tier coefficients. Since comparison to the individual condition model is the main point of this table and the analog for this method for the individual condition model is not clear, we postpone describing this method for handling insignificant coefficients until we discuss weights in Section 6.

conditions model to produce Models 7, 8, and 9.[10] (Model 3 would be essentially unchanged because it would drop the comorbidity payment only for the handful of cases with comorbidity in RIC 21; Model 2 is not a reasonable candidate because it omits conditions that we know affect cost.)

After dropping insignificant ($p > 0.05$) coefficients, the ranking of the three models remains the same: The individual conditions model explains the most variance, followed by the "index model" based on condition counts, followed by the single-tier-assignment model. However, the single-tier assignment model achieves 85.8 percent of the possible increase in R-squared from the no-comorbidity model. Counting all conditions within each tier results in 91.6 percent of the maximum R-squared. Thus, all conditions within a single tier have very similar cost effects, and we lose little accuracy by assuming that they are exactly the same.

Table 4.8 shows the comorbidity coefficients in each RIC for models 2, 3, 4, and 5. As we found in our preliminary analyses for the NPRM, the comorbidity variable that we recommended at that time (Model 2) is not statistically significant in small RICs (11, 18, 19, and 21 in the 1998-1999 data).

The additional comorbidities that are accounted for in our new work result in statistical significance of a single indicator variable (Model 3) in all RICs except the tiny burn RIC (21). There are only 38 cases with a relevant comorbidity in this RIC.

As may be seen by comparing the coefficient on the comorbidity variable in Model 3 with the coefficients on the three tier variables in the same RIC, the single comorbidity variable misses substantial variation in the cost of different comorbidities. For example, in RIC 1, although the average case with any comorbidity costs only about 5 percent more than cases with no comorbidity, cases with a tier 1 comorbidity average more

---

[10] The variables dropped in Models 7 and 8 can be determined from Table 4.8 as those that are negative or not significant.

Table 4.8

Coefficients on Comorbidity Variables in Selected Regressions of Log (Cost) on FRG and Comorbidity Variables (pooled 1998-1999 data)

| Model | 2. Single Comorbidity as Defined in NPRM | | 3. Single Comorbidity, Table 2 Conditions | | 4. Tier Assignment | | | | | | 5. Count Conditions | | | | | |
| | | | | | Tier 1 | | Tier 2 | | Tier 3 | | Tier 1 | | Tier 2 | | Tier 3 | |
RIC	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.
1	0.0484	7.28	0.0483	9.15	0.1867	7.83	0.0760	6.56	0.0342	5.78	0.1719	7.45	0.0656	5.90	0.0291	5.93
2	0.0835	3.96	0.1408	8.41	0.1975	3.60	0.1920	7.97	0.1006	4.86	0.1215	2.27	0.1550	7.08	0.0802	5.07
3	0.0963	6.44	0.1087	9.19	0.2297	4.63	0.1418	8.07	0.0854	6.22	0.1371	2.94	0.0999	6.37	0.0667	6.37
4	0.1699	4.52	0.1585	5.78	0.3594	4.69	0.1603	3.89	0.1251	3.45	0.2342	3.54	0.1423	3.96	0.1182	4.05
5	0.2142	11.70	0.1904	15.38	0.3631	6.09	0.2543	13.68	0.1399	8.89	0.2501	4.47	0.1910	11.20	0.1134	8.91
6	0.0897	7.68	0.1022	11.67	0.3281	6.76	0.1094	8.63	0.0889	8.07	0.2324	5.20	0.0853	7.40	0.0735	8.90
7	0.1453	12.39	0.1269	16.37	0.0726[a]	1.50	0.1531	12.01	0.1158	12.18	0.0383	0.79	0.1244	10.37	0.1037	13.11
8	0.1823	19.51	0.1703	27.29	0.3123	6.59	0.1910	20.10	0.1523	18.86	0.2792	5.89	0.1664	18.25	0.1373	19.2
9	0.1962	12.00	0.1611	14.71	0.3216	4.58	0.1906	10.32	0.1431	10.87	0.2565	3.65	0.1527	8.66	0.1197	10.83
10	0.0719	6.39	0.0973	10.9	0.1853	2.94	0.1535	9.52	0.0820	8.50	0.1244	2.03	0.1056	7.28	0.0735	10.76
11	0.0396[a]	1.12	0.1096	3.96	0.6138	2.20	0.1570	3.64	0.0938	3.17	0.5377[a]	1.93	0.0778	2.14	0.0348[a]	1.87
12	0.1477	6.66	0.1587	10.32	0.5994	5.56	0.1974	7.58	0.1288	6.98	0.5179	4.80	0.1673	6.74	0.1288	8.15
13	0.0926	3.41	0.1486	7.08	0.3653	2.62	0.1658	4.95	0.1353	5.24	0.2891	2.07	0.1157	3.61	0.1066	5.16
14	0.1343	11.82	0.1349	14.8	0.3194	8.38	0.2052	14.97	0.0799	7.16	0.2226	6.21	0.1630	13.09	0.0623	7.54
15	0.0716	5.45	0.0904	6.49	0.2302	2.86	0.1466	7.91	0.0201[a]	1.02	0.1752	2.27	0.1331	7.82	0.0136[a]	0.86
16	0.2163	7.02	0.1852	9.59	0.2575[a]	1.93	0.2075	6.54	0.1727	7.37	0.1422[a]	1.06	0.1581	5.28	0.1432	7.40
17	0.1459	4.64	0.1365	5.69	0.2628	3.59	0.2119	5.26	0.0779	2.58	0.1961	2.75	0.1848	5.18	0.0730	2.94
18	0.0201[a]	-0.27	0.1038[a]	1.96	0.1458[a]	1.21	0.2209	2.94	-0.029[a]	-0.39	0.0748[a]	0.62	0.2221	3.38	-0.048[a]	-0.88
19	0.0766[a]	1.30	0.0986	2.25	0.0811[a]	0.66	0.1462	2.47	0.0555[a]	0.92	0.0295[a]	0.26	0.1500	2.85	0.0815	1.75
20	0.0942	15.46	0.1211	23.67	0.3016	15.01	0.1572	20.66	0.0909	14.93	0.2138	11.58	0.1109	16.60	0.0710	16.38
21	-0.0457[a]	-0.50	-0.051[a]	-0.7	0.0308[a]	0.13	0.0260[a]	0.23	-0.099[a]	-1.12	0.0393[a]	0.16	0.0673[a]	0.74	-0.026[a]	-0.4
R-sq	0.3352		0.3379		0.3388						0.3392					

[a] $p > 0.05$.

Note: Each regression controlled for each FRG and for year.

than 18 percent more than cases with no comorbidity. Even larger coefficients are found in many RICs. In Section 3, we discussed our decision not to combine FRGs when it would change estimated cost by $1,000 or more, or 8.9 percent of the average cost ($11,199 in this sample). In almost all RICs the difference between the tier 1 coefficient and the Model 3 coefficient is greater than this percentage.

As one can see by examining coefficients for the same tier across RICs, differences across RICs in the effect of comorbidities are also quite large. In dollar terms, these differences make it easy enough to justify payment differences. In 17 of the 21 RICs (all except RICs 7, 18, 19, and 21), the coefficients are monotonic in tier--i.e., tier 1 cases cost more than tier 2 cases, which cost more than tier 3 cases, which cost more than cases with no comorbidities. The violations of monotonicity usually reverse the tier 1 and tier 2, and we believe that this arises from random variation in the cost of the very infrequent tier 1 cases. Tier 1 conditions are very rare in the orthopedic RICs, so even the large RIC 7 has only 50 tier 1 cases in the sample. Consequently, we believe it reasonable to define tiers using analyses of the marginal effect of individual comorbidities across RICs, but then to use RIC-specific data to measure the relative cost of each tier.

The last columns of Table 4.8 (Model 5) show the coefficients on the count of the number of conditions that the patient has in each tier. In Model 4, each tier coefficient covers the cost of all conditions that the patient has in that tier and in all lower tiers. Thus the difference in the tier 1 coefficients between Models 4 and 5 is the expected cost of additional secondary conditions in persons who have at least one tier 1 condition. The difference in tier 2 coefficients is the expected cost of additional secondary conditions in persons who have at least one tier 2 condition but no tier 1 conditions. And the difference in tier 3 coefficients is the expected cost of additional secondary conditions in persons who have at least one tier 3 condition but no tier 1 or 2 conditions. In almost all RICs, the differences between Models 4 and 5 in tier 2 and tier 3 coefficients are quite small and not even statistically

significant. However, the difference in tier 1 coefficients is sometimes quite large (see, e.g., RICs 3, 4, 5).

For completeness, Table 4.9 shows the comorbidity coefficients for Models 2 through 5 when they are fit on the 1996-1997 data. A comparison of Tables 4.8 and 4.9 shows the relative stability of the models. The single comorbidity variables are the most stable: The average absolute value of the change in the 21 coefficients is about 0.04. The change in the absolute value of the tier 2 and tier 3 coefficients also is typically about 0.04. However, the tier 1 coefficients change by an average of 0.14 in both Models 4 and 5. Unfortunately, we cannot say whether this instability is due to randomness, changed coding practices, or the apparent real trends we saw in Table 4.2 in the cost of care for ventilator and tracheostomy patients.

While the stability of the tier 1 coefficients is not all that one would hope for, the stability of all the coefficients in Models 4 and 5 greatly exceeds that of individual conditions interacted by RIC. Although we do not show the details, the average change in absolute value of the comorbidity coefficients of the model was 0.17--with diseases in tier 1 changing by 0.23, those in tier 2 by 0.17, and those in tier 1 by 0.13.

As a final check on the stability of the tiered model, we used a tenfold cross-validation model to verify that our model of predicting cost from FRG and the interaction of tier and RIC was not seriously affected by sampling issues. The RMSE increased by only 0.13 percent (i.e., a little over one-tenth of 1 percent) in the out-of-sample prediction, from the value in the fitted sample.

Table 4.9

Coefficients on Comorbidity Variables in Selected Regressions of Log (Cost) on FRG and Comorbidity Variables
(pooled 1996-1997 data)

Model	2. Single Comorbidity as Defined in NPRM		3. Single Comorbidity, Table 2 Conditions		4. Tier Assignment						5. Count Conditions					
					Tier 1		Tier 2		Tier 3		Tier 1		Tier 2		Tier 3	
RIC	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.
1	0.0521	6.67	0.0430	7.76	0.1385	5.67	0.0758	6.14	0.0298	4.81	0.1313	5.52	0.0690	5.84	0.0280	5.38
2	0.0200[a]	0.84	0.0910	4.92	0.1952	2.78	0.1744	6.34	0.0348[a]	1.57	0.1546	2.29	0.1534	5.97	0.0385	2.23
3	0.0922	5.57	0.1084	8.31	0.2903	5.39	0.1248	6.11	0.0937	6.42	0.1904	4.10	0.0761	4.15	0.0704	6.36
4	0.2077	5.10	0.1525	5.16	0.4312	4.84	0.2331	5.13	0.0489[a]	1.28	0.3306	4.26	0.2043	4.86	0.0670	2.15
5	0.1941	9.63	0.1552	11.53	0.4205	7.09	0.2095	10.04	0.1080	6.47	0.3300	5.69	0.1616	8.35	0.0857	6.40
6	0.0733	5.06	0.0742	6.99	0.2481	5.33	0.0952	6.13	0.0483	3.63	0.2104	4.71	0.0737	5.24	0.0482	4.96
7	0.1330	10.90	0.1118	14.65	0.1151	2.66	0.1440	10.40	0.0996	11.18	0.0862	2.02	0.1111	8.43	0.0889	12.18
8	0.1878	18.24	0.1585	23.98	0.2315	5.00	0.1946	18.65	0.1341	16.07	0.1816	4.09	0.1574	15.61	0.1221	17.20
9	0.1893	10.06	0.1456	11.73	0.2640	3.58	0.2190	9.98	0.1120	7.69	0.1673	2.42	0.1701	8.45	0.0933	7.96
10	0.0773	6.36	0.0904	9.57	0.2183	3.09	0.1791	9.78	0.0696	6.92	0.1811	2.82	0.1345	7.95	0.0574	8.11
11	0.1092	2.75	0.1631	4.94	0.1399[a]	0.66	0.2806	5.09	0.1338	3.86	0.0619[a]	0.29	0.1578	3.57	0.0721	3.46
12	0.1250	4.69	0.1751	9.46	0.2837	2.25	0.1515	4.41	0.1813	8.53	0.2318[a]	1.84	0.0954	2.93	0.1575	8.87
13	0.1352	4.05	0.1375	5.58	-0.0442[a]	-0.28	0.1933	4.90	0.1142	3.81	-0.1327[a]	-0.84	0.1496	4.00	0.0974	3.93
14	0.1685	12.01	0.1480	13.74	0.3829	8.14	0.2082	12.55	0.1030	8.00	0.2919	6.48	0.1433	9.40	0.0900	9.63
15	0.0452	2.43	0.0462	2.90	0.1695	2.60	0.1174	5.42	-0.041[a]	-1.87	0.1514	2.39	0.1184	6.11	0.0191[a]	-1.13
16	0.2091	5.56	0.1297	5.49	0.3336	2.12	0.1435	3.53	0.1181	4.22	0.2104[a]	1.72	0.0869	2.20	0.1233	5.30
17	0.2182	5.78	0.1433	5.40	0.4346	5.16	0.2493	5.33	0.0566[a]	1.74	0.3610	4.64	0.1791	4.39	0.0615	2.29
18	0.2404	3.51	0.2004	3.72	0.5514	4.41	0.2947	3.54	0.0461[a]	0.66	0.4333	3.44	0.2407	3.37	0.0376[a]	0.71
19	0.0910[a]	1.26	0.0806[a]	1.64	0.2888	2.38	0.1566	2.32	-0.067[a]	-0.93	0.2372	2.13	0.1562	2.50	0.0073[a]	-0.15
20	0.0931	12.25	0.1041	16.40	0.3197	12.86	0.1513	15.62	0.0684	9.32	0.2508	10.57	0.1125	13.49	0.0595	11.77
21	0.0824[a]	0.90	0.2372	3.28	0.3020[a]	1.40	0.2663	2.36	0.2206	2.62	0.1938[a]	0.91	0.2247	2.09	0.0978[a]	1.82
R-sq	0.3352		0.3551		0.3562						0.3568					

[a] $p > 0.05$.

Note: Each regression controlled for each FRG and for year.

CONCLUSIONS

We have seen that cases with certain comorbidities cost substantially more than other cases in the same FRG. Further, the list of comorbidities that demonstrably increase cost is much larger than the list of comorbidities found in the NPRM (which was based on our preliminary recommendation to HCFA).

For the final rule, we recommended using three tiers to pay for relevant comorbidities in the initial year(s) of the IRF PPS. We found that some infrequent comorbidities cost substantially more than other comorbidities, and we believe the difference is large enough to require payment to avoid problems with access and to provide adequate resources for these expensive patients. The facility would receive a payment according to the most expensive tier for which the patient qualifies. Because of problems with ICD-9-CM coding in our database, we believe that paying a separate amount for each comorbidity might seriously threaten budget neutrality and could conceivably lead to upcoding. Further, paying a different amount for each comorbidity would produce very unstable payments and add very little to the accuracy of the cost prediction.

We recommended that the tiers be defined consistently across each RIC (except for RIC-specific exclusions) but that the amount of the additional payment for each tier be calculated based on RIC-specific data. The effects of RIC are statistically important and have been consistent in every model we tried. Within each RIC, cost effects consistently vary with tiers defined based on their overall effect.

Although we are confident that this recommendation will lead to reasonable payment amounts, there are several ways in which we would like to pursue refinement in the coming years when we will have a more representative and better coded database. First, we would like to systematically address all ICD-9-CM codes, rather than just the ones hypothesized by our clinical experts. The cancer analysis above provides a model that could be applied to all the codes in each body system within the ICD-9-CM coding system. As part of this exploration we would like to separately address the comorbidities that are present at admission. It may be that the comorbidities that were removed from the payment list because

they may be preventable have substantial costs when present at admission
and should be reimbursed.

Second, we would like to pursue the possibility of a complexity
adjustment that better takes into account the cost of cases with multiple
comorbidities (as well as multiple impairments). We have already seen in
the count conditions model that an important fraction of the cost of tier
1 conditions is due to multiple comorbidities. We hope to be able to
develop a complexity model which accounts for these interactions without
having the incentives for upcoding that are found in a simple count model.
Many medically complex cases are found in RIC 20, the miscellaneous RIC.
By detailing the clinical characteristics of all patients in RIC 20, we
may be able to discover sub-groups that should be treated as separate
RICs.

Third, we would like to integrate the analysis of comorbidities with
the creation of FRGs to develop a single algorithm that produces FRGCs.
Because comorbidities are more frequent in cases with very low function,
it may be possible to produce better groups if we account for potential
comorbidity payment at the time of creation of the FRGCs.

Our analysis also has implications for monitoring. In particular,
the apparent trend in the tier 1 coefficients suggests the possibility
that we have overestimated the cost of these infrequent conditions by
using 1998-1999 data. Further, the rate of coding of comorbidities may
increase and threaten budget neutrality. Thus, comorbidity coding and the
costs of comorbidity cases should be carefully watched in the initial year
of the rehabilitation PPS.

Table 4.10

Recommendation for Additional Payment:
ICD-9-CM Diagnoses, Tier, and RIC Exclusions

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
008.42	PSEUDOMONAS ENTERITIS	2		Pseudomonas
008.45	INT INF CLSTRDIUM DFCILE	2		Intestinal infection clostridium
011.	PULMONARY TUBERCULOSIS*	2	15	Other infections
011.0	TB OF LUNG_ INFILTRATIVE*	2	15	Other infections
011.00	TB LUNG INFILTR-UNSPEC	2	15	Other infections
011.01	TB LUNG INFILTR-NO EXAM	2	15	Other infections
011.02	TB LUNG INFILTR-EXM UNKN	2	15	Other infections
011.03	TB LUNG INFILTR-MICRO DX	2	15	Other infections
011.04	TB LUNG INFILTR-CULT DX	2	15	Other infections
011.05	TB LUNG INFILTR-HISTO DX	2	15	Other infections
011.06	TB LUNG INFILTR-OTH TEST	2	15	Other infections
011.1	TB OF LUNG_ NODULAR*	2	15	Other infections
011.10	TB LUNG NODULAR-UNSPEC	2	15	Other infections
011.11	TB LUNG NODULAR-NO EXAM	2	15	Other infections
011.12	TB LUNG NODUL-EXAM UNKN	2	15	Other infections
011.13	TB LUNG NODULAR-MICRO DX	2	15	Other infections
011.14	TB LUNG NODULAR-CULT DX	2	15	Other infections
011.15	TB LUNG NODULAR-HISTO DX	2	15	Other infections
011.16	TB LUNG NODULAR-OTH TEST	2	15	Other infections
011.2	TB OF LUNG W CAVITATION*	2	15	Other infections
011.20	TB LUNG W CAVITY-UNSPEC	2	15	Other infections
011.21	TB LUNG W CAVITY-NO EXAM	2	15	Other infections
011.22	TB LUNG CAVITY-EXAM UNKN	2	15	Other infections
011.23	TB LUNG W CAVIT-MICRO DX	2	15	Other infections
011.24	TB LUNG W CAVITY-CULT DX	2	15	Other infections
011.25	TB LUNG W CAVIT-HISTO DX	2	15	Other infections
011.26	TB LUNG W CAVIT-OTH TEST	2	15	Other infections
011.3	TUBERCULOSIS OF BRONCHUS*	2	15	Other infections
011.30	TB OF BRONCHUS-UNSPEC	2	15	Other infections
011.31	TB OF BRONCHUS-NO EXAM	2	15	Other infections
011.32	TB OF BRONCHUS-EXAM UNKN	2	15	Other infections
011.33	TB OF BRONCHUS-MICRO DX	2	15	Other infections
011.34	TB OF BRONCHUS-CULT DX	2	15	Other infections
011.35	TB OF BRONCHUS-HISTO DX	2	15	Other infections
011.36	TB OF BRONCHUS-OTH TEST	2	15	Other infections
011.4	TB FIBROSIS OF LUNG*	2	15	Other infections
011.40	TB LUNG FIBROSIS-UNSPEC	2	15	Other infections
011.41	TB LUNG FIBROSIS-NO EXAM	2	15	Other infections

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
011.42	TB LUNG FIBROS-EXAM UNKN	2	15	Other infections
011.43	TB LUNG FIBROS-MICRO DX	2	15	Other infections
011.44	TB LUNG FIBROSIS-CULT DX	2	15	Other infections
011.45	TB LUNG FIBROS-HISTO DX	2	15	Other infections
011.46	TB LUNG FIBROS-OTH TEST	2	15	Other infections
011.5	TB BRONCHIECTASIS*	2	15	Other infections
011.50	TB BRONCHIECTASIS-UNSPEC	2	15	Other infections
011.51	TB BRONCHIECT-NO EXAM	2	15	Other infections
011.52	TB BRONCHIECT-EXAM UNKN	2	15	Other infections
011.53	TB BRONCHIECT-MICRO DX	2	15	Other infections
011.54	TB BRONCHIECT-CULT DX	2	15	Other infections
011.55	TB BRONCHIECT-HISTO DX	2	15	Other infections
011.56	TB BRONCHIECT-OTH TEST	2	15	Other infections
011.6	TUBERCULOUS PNEUMONIA*	2	15	Other infections
011.60	TB PNEUMONIA-UNSPEC	2	15	Other infections
011.61	TB PNEUMONIA-NO EXAM	2	15	Other infections
011.62	TB PNEUMONIA-EXAM UNKN	2	15	Other infections
011.63	TB PNEUMONIA-MICRO DX	2	15	Other infections
011.64	TB PNEUMONIA-CULT DX	2	15	Other infections
011.65	TB PNEUMONIA-HISTO DX	2	15	Other infections
011.66	TB PNEUMONIA-OTH TEST	2	15	Other infections
011.7	TUBERCULOUS PNEUMOTHORAX*	2	15	Other infections
011.70	TB PNEUMOTHORAX-UNSPEC	2	15	Other infections
011.71	TB PNEUMOTHORAX-NO EXAM	2	15	Other infections
011.72	TB PNEUMOTHORX-EXAM UNKN	2	15	Other infections
011.73	TB PNEUMOTHORAX-MICRO DX	2	15	Other infections
011.74	TB PNEUMOTHORAX-CULT DX	2	15	Other infections
011.75	TB PNEUMOTHORAX-HISTO DX	2	15	Other infections
011.76	TB PNEUMOTHORAX-OTH TEST	2	15	Other infections
011.8	PULMONARY TB NEC*	2	15	Other infections
011.80	PULMONARY TB NEC-UNSPEC	2	15	Other infections
011.81	PULMONARY TB NEC-NO EXAM	2	15	Other infections
011.82	PULMON TB NEC-EXAM UNKN	2	15	Other infections
011.83	PULMON TB NEC-MICRO DX	2	15	Other infections
011.84	PULMON TB NEC-CULT DX	2	15	Other infections
011.85	PULMON TB NEC-HISTO DX	2	15	Other infections
011.86	PULMON TB NEC-OTH TEST	2	15	Other infections
011.9	PULMONARY TB NOS*	2	15	Other infections
011.90	PULMONARY TB NOS-UNSPEC	2	15	Other infections
011.91	PULMONARY TB NOS-NO EXAM	2	15	Other infections
011.92	PULMON TB NOS-EXAM UNKN	2	15	Other infections
011.93	PULMON TB NOS-MICRO DX	2	15	Other infections
011.94	PULMON TB NOS-CULT DX	2	15	Other infections
011.95	PULMON TB NOS-HISTO DX	2	15	Other infections
011.96	PULMON TB NOS-OTH TEST	2	15	Other infections

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
012.	OTHER RESPIRATORY TB*	2	15	Other infections
012.0	TUBERCULOUS PLEURISY*	2	15	Other infections
012.00	TB PLEURISY-UNSPEC	2	15	Other infections
012.01	TB PLEURISY-NO EXAM	2	15	Other infections
012.02	TB PLEURISY-EXAM UNKN	2	15	Other infections
012.03	TB PLEURISY-MICRO DX	2	15	Other infections
012.04	TB PLEURISY-CULT DX	2	15	Other infections
012.05	TB PLEURISY-HISTOLOG DX	2	15	Other infections
012.06	TB PLEURISY-OTH TEST	2	15	Other infections
012.1	TB THORACIC LYMPH NODES*	2	15	Other infections
012.10	TB THORACIC NODES-UNSPEC	2	15	Other infections
012.11	TB THORAX NODE-NO EXAM	2	15	Other infections
012.12	TB THORAX NODE-EXAM UNKN	2	15	Other infections
012.13	TB THORAX NODE-MICRO DX	2	15	Other infections
012.14	TB THORAX NODE-CULT DX	2	15	Other infections
012.15	TB THORAX NODE-HISTO DX	2	15	Other infections
012.16	TB THORAX NODE-OTH TEST	2	15	Other infections
012.2	ISOLATED TRACH/BRONCH TB*	2	15	Other infections
012.20	ISOL TRACHEAL TB-UNSPEC	2	15	Other infections
012.21	ISOL TRACHEAL TB-NO EXAM	2	15	Other infections
012.22	ISOL TRACH TB-EXAM UNKN	2	15	Other infections
012.23	ISOLAT TRACH TB-MICRO DX	2	15	Other infections
012.24	ISOL TRACHEAL TB-CULT DX	2	15	Other infections
012.25	ISOLAT TRACH TB-HISTO DX	2	15	Other infections
012.26	ISOLAT TRACH TB-OTH TEST	2	15	Other infections
012.3	TUBERCULOUS LARYNGITIS*	2	15	Other infections
012.30	TB LARYNGITIS-UNSPEC	2	15	Other infections
012.31	TB LARYNGITIS-NO EXAM	2	15	Other infections
012.32	TB LARYNGITIS-EXAM UNKN	2	15	Other infections
012.33	TB LARYNGITIS-MICRO DX	2	15	Other infections
012.34	TB LARYNGITIS-CULT DX	2	15	Other infections
012.35	TB LARYNGITIS-HISTO DX	2	15	Other infections
012.36	TB LARYNGITIS-OTH TEST	2	15	Other infections
012.8	RESPIRATORY TB NEC*	2	15	Other infections
012.80	RESP TB NEC-UNSPEC	2	15	Other infections
012.81	RESP TB NEC-NO EXAM	2	15	Other infections
012.82	RESP TB NEC-EXAM UNKN	2	15	Other infections
012.83	RESP TB NEC-MICRO DX	2	15	Other infections
012.84	RESP TB NEC-CULT DX	2	15	Other infections
012.85	RESP TB NEC-HISTO DX	2	15	Other infections
012.86	RESP TB NEC-OTH TEST	2	15	Other infections
013.	CNS TUBERCULOSIS*	2	3,5	Other infections
013.0	TUBERCULOUS MENINGITIS*	2	3,5	Other infections
013.00	TB MENINGITIS-UNSPEC	2	3,5	Other infections
013.01	TB MENINGITIS-NO EXAM	2	3,5	Other infections

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
013.02	TB MENINGITIS-EXAM UNKN	2	3,5	Other infections
013.03	TB MENINGITIS-MICRO DX	2	3,5	Other infections
013.04	TB MENINGITIS-CULT DX	2	3,5	Other infections
013.05	TB MENINGITIS-HISTO DX	2	3,5	Other infections
013.06	TB MENINGITIS-OTH TEST	2	3,5	Other infections
013.1	TUBERCULOMA OF MENINGES*	2	3,5	Other infections
013.10	TUBRCLMA MENINGES-UNSPEC	2	3,5	Other infections
013.11	TUBRCLMA MENING-NO EXAM	2	3,5	Other infections
013.12	TUBRCLMA MENIN-EXAM UNKN	2	3,5	Other infections
013.13	TUBRCLMA MENING-MICRO DX	2	3,5	Other infections
013.14	TUBRCLMA MENING-CULT DX	2	3,5	Other infections
013.15	TUBRCLMA MENING-HISTO DX	2	3,5	Other infections
013.16	TUBRCLMA MENING-OTH TEST	2	3,5	Other infections
013.2	TUBERCULOMA OF BRAIN*	2	3	Other infections
013.20	TUBERCULOMA BRAIN-UNSPEC	2	3	Other infections
013.21	TUBRCLOMA BRAIN-NO EXAM	2	3	Other infections
013.22	TUBRCLMA BRAIN-EXAM UNKN	2	3	Other infections
013.23	TUBRCLOMA BRAIN-MICRO DX	2	3	Other infections
013.24	TUBRCLOMA BRAIN-CULT DX	2	3	Other infections
013.25	TUBRCLOMA BRAIN-HISTO DX	2	3	Other infections
013.26	TUBRCLOMA BRAIN-OTH TEST	2	3	Other infections
013.3	TB ABSCESS OF BRAIN*	2	3	Other infections
013.30	TB BRAIN ABSCESS-UNSPEC	2	3	Other infections
013.31	TB BRAIN ABSCESS-NO EXAM	2	3	Other infections
013.32	TB BRAIN ABSC-EXAM UNKN	2	3	Other infections
013.33	TB BRAIN ABSC-MICRO DX	2	3	Other infections
013.34	TB BRAIN ABSCESS-CULT DX	2	3	Other infections
013.35	TB BRAIN ABSC-HISTO DX	2	3	Other infections
013.36	TB BRAIN ABSC-OTH TEST	2	3	Other infections
013.4	TUBERCULOMA SPINAL CORD*	2	5	Other infections
013.40	TUBRCLMA SP CORD-UNSPEC	2	5	Other infections
013.41	TUBRCLMA SP CORD-NO EXAM	2	5	Other infections
013.42	TUBRCLMA SP CD-EXAM UNKN	2	5	Other infections
013.43	TUBRCLMA SP CRD-MICRO DX	2	5	Other infections
013.44	TUBRCLMA SP CORD-CULT DX	2	5	Other infections
013.45	TUBRCLMA SP CRD-HISTO DX	2	5	Other infections
013.46	TUBRCLMA SP CRD-OTH TEST	2	5	Other infections
013.5	TB ABSCESS SPINAL CORD*	2	5	Other infections
013.50	TB SP CRD ABSCESS-UNSPEC	2	5	Other infections
013.51	TB SP CRD ABSC-NO EXAM	2	5	Other infections
013.52	TB SP CRD ABSC-EXAM UNKN	2	5	Other infections
013.53	TB SP CRD ABSC-MICRO DX	2	5	Other infections
013.54	TB SP CRD ABSC-CULT DX	2	5	Other infections
013.55	TB SP CRD ABSC-HISTO DX	2	5	Other infections
013.56	TB SP CRD ABSC-OTH TEST	2	5	Other infections

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
013.6	TB ENCEPHALITIS/MYELITIS*	2	3	Other infections
013.60	TB ENCEPHALITIS-UNSPEC	2	3	Other infections
013.61	TB ENCEPHALITIS-NO EXAM	2	3	Other infections
013.62	TB ENCEPHALIT-EXAM UNKN	2	3	Other infections
013.63	TB ENCEPHALITIS-MICRO DX	2	3	Other infections
013.64	TB ENCEPHALITIS-CULT DX	2	3	Other infections
013.65	TB ENCEPHALITIS-HISTO DX	2	3	Other infections
013.66	TB ENCEPHALITIS-OTH TEST	2	3	Other infections
013.8	CNS TUBERCULOSIS NEC*	2	3,5	Other infections
013.80	CNS TB NEC-UNSPEC	2	3,5	Other infections
013.81	CNS TB NEC-NO EXAM	2	3,5	Other infections
013.82	CNS TB NEC-EXAM UNKN	2	3,5	Other infections
013.83	CNS TB NEC-MICRO DX	2	3,5	Other infections
013.84	CNS TB NEC-CULT DX	2	3,5	Other infections
013.85	CNS TB NEC-HISTO DX	2	3,5	Other infections
013.86	CNS TB NEC-OTH TEST	2	3,5	Other infections
013.9	CNS TUBERCULOSIS NOS*	2	3,5	Other infections
013.90	CNS TB NOS-UNSPEC	2	3,5	Other infections
013.91	CNS TB NOS-NO EXAM	2	3,5	Other infections
013.92	CNS TB NOS-EXAM UNKN	2	3,5	Other infections
013.93	CNS TB NOS-MICRO DX	2	3,5	Other infections
013.94	CNS TB NOS-CULT DX	2	3,5	Other infections
013.95	CNS TB NOS-HISTO DX	2	3,5	Other infections
013.96	CNS TB NOS-OTH TEST	2	3,5	Other infections
014.	INTESTINAL TB*	2		Other infections
014.0	TUBERCULOUS PERITONITIS*	2		Other infections
014.00	TB PERITONITIS-UNSPEC	2		Other infections
014.01	TB PERITONITIS-NO EXAM	2		Other infections
014.02	TB PERITONITIS-EXAM UNKN	2		Other infections
014.03	TB PERITONITIS-MICRO DX	2		Other infections
014.04	TB PERITONITIS-CULT DX	2		Other infections
014.05	TB PERITONITIS-HISTO DX	2		Other infections
014.06	TB PERITONITIS-OTH TEST	2		Other infections
014.8	INTESTINAL TB NEC*	2		Other infections
014.80	INTESTINAL TB NEC-UNSPEC	2		Other infections
014.81	INTESTIN TB NEC-NO EXAM	2		Other infections
014.82	INTEST TB NEC-EXAM UNKN	2		Other infections
014.83	INTESTIN TB NEC-MICRO DX	2		Other infections
014.84	INTESTIN TB NEC-CULT DX	2		Other infections
014.85	INTESTIN TB NEC-HISTO DX	2		Other infections
014.86	INTESTIN TB NEC-OTH TEST	2		Other infections
015.	TB OF BONE AND JOINT*	2	3,9	Other infections
015.0	TB OF VERTEBRAL COLUMN*	2	3,9	Other infections
015.00	TB OF VERTEBRA-UNSPEC	2	3,9	Other infections
015.01	TB OF VERTEBRA-NO EXAM	2	3,9	Other infections

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
015.02	TB OF VERTEBRA-EXAM UNKN	2	3,9	Other infections
015.03	TB OF VERTEBRA-MICRO DX	2	3,9	Other infections
015.04	TB OF VERTEBRA-CULT DX	2	3,9	Other infections
015.05	TB OF VERTEBRA-HISTO DX	2	3,9	Other infections
015.06	TB OF VERTEBRA-OTH TEST	2	3,9	Other infections
015.1	TB OF HIP*	2	9	Other infections
015.10	TB OF HIP-UNSPEC	2	9	Other infections
015.11	TB OF HIP-NO EXAM	2	9	Other infections
015.12	TB OF HIP-EXAM UNKN	2	9	Other infections
015.13	TB OF HIP-MICRO DX	2	9	Other infections
015.14	TB OF HIP-CULT DX	2	9	Other infections
015.15	TB OF HIP-HISTO DX	2	9	Other infections
015.16	TB OF HIP-OTH TEST	2	9	Other infections
015.2	TB OF KNEE*	2	9	Other infections
015.20	TB OF KNEE-UNSPEC	2	9	Other infections
015.21	TB OF KNEE-NO EXAM	2	9	Other infections
015.22	TB OF KNEE-EXAM UNKN	2	9	Other infections
015.23	TB OF KNEE-MICRO DX	2	9	Other infections
015.24	TB OF KNEE-CULT DX	2	9	Other infections
015.25	TB OF KNEE-HISTO DX	2	9	Other infections
015.26	TB OF KNEE-OTH TEST	2	9	Other infections
015.5	TB OF LIMB BONES*	2	9,10,11	Other infections
015.50	TB OF LIMB BONES-UNSPEC	2	9,10,11	Other infections
015.51	TB LIMB BONES-NO EXAM	2	9,10,11	Other infections
015.52	TB LIMB BONES-EXAM UNKN	2	9,10,11	Other infections
015.53	TB LIMB BONES-MICRO DX	2	9,10,11	Other infections
015.54	TB LIMB BONES-CULT DX	2	9,10,11	Other infections
015.55	TB LIMB BONES-HISTO DX	2	9,10,11	Other infections
015.56	TB LIMB BONES-OTH TEST	2		Other infections
015.6	TB OF MASTOID*	2		Other infections
015.60	TB OF MASTOID-UNSPEC	2		Other infections
015.61	TB OF MASTOID-NO EXAM	2		Other infections
015.62	TB OF MASTOID-EXAM UNKN	2		Other infections
015.63	TB OF MASTOID-MICRO DX	2		Other infections
015.64	TB OF MASTOID-CULT DX	2		Other infections
015.65	TB OF MASTOID-HISTO DX	2		Other infections
015.66	TB OF MASTOID-OTH TEST	2		Other infections
015.7	TB OF BONE NEC*	2	9	Other infections
015.70	TB OF BONE NEC-UNSPEC	2	9	Other infections
015.71	TB OF BONE NEC-NO EXAM	2	9	Other infections
015.72	TB OF BONE NEC-EXAM UNKN	2	9	Other infections
015.73	TB OF BONE NEC-MICRO DX	2	9	Other infections
015.74	TB OF BONE NEC-CULT DX	2	9	Other infections
015.75	TB OF BONE NEC-HISTO DX	2	9	Other infections
015.76	TB OF BONE NEC-OTH TEST	2	9	Other infections

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
015.8	TB OF JOINT NEC*	2	9	Other infections
015.80	TB OF JOINT NEC-UNSPEC	2	9	Other infections
015.81	TB OF JOINT NEC-NO EXAM	2	9	Other infections
015.82	TB JOINT NEC-EXAM UNKN	2	9	Other infections
015.83	TB OF JOINT NEC-MICRO DX	2	9	Other infections
015.84	TB OF JOINT NEC-CULT DX	2	9	Other infections
015.85	TB OF JOINT NEC-HISTO DX	2	9	Other infections
015.86	TB OF JOINT NEC-OTH TEST	2	9	Other infections
015.9	TB OF BONE & JOINT NOS*	2	9	Other infections
015.90	TB BONE/JOINT NOS-UNSPEC	2	9	Other infections
015.91	TB BONE/JT NOS-NO EXAM	2	9	Other infections
015.92	TB BONE/JT NOS-EXAM UNKN	2	9	Other infections
015.93	TB BONE/JT NOS-MICRO DX	2	9	Other infections
015.94	TB BONE/JT NOS-CULT DX	2	9	Other infections
015.95	TB BONE/JT NOS-HISTO DX	2	9	Other infections
015.96	TB BONE/JT NOS-OTH TEST	2	9	Other infections
016.	GENITOURINARY TB*	2		Other infections
016.0	TB OF KIDNEY*	2		Other infections
016.00	TB OF KIDNEY-UNSPEC	2		Other infections
016.01	TB OF KIDNEY-NO EXAM	2		Other infections
016.02	TB OF KIDNEY-EXAM UNKN	2		Other infections
016.03	TB OF KIDNEY-MICRO DX	2		Other infections
016.04	TB OF KIDNEY-CULT DX	2		Other infections
016.05	TB OF KIDNEY-HISTO DX	2		Other infections
016.06	TB OF KIDNEY-OTH TEST	2		Other infections
016.1	TB OF BLADDER*	2		Other infections
016.10	TB OF BLADDER-UNSPEC	2		Other infections
016.11	TB OF BLADDER-NO EXAM	2		Other infections
016.12	TB OF BLADDER-EXAM UNKN	2		Other infections
016.13	TB OF BLADDER-MICRO DX	2		Other infections
016.14	TB OF BLADDER-CULT DX	2		Other infections
016.15	TB OF BLADDER-HISTO DX	2		Other infections
016.16	TB OF BLADDER-OTH TEST	2		Other infections
016.2	TB OF URETER*	2		Other infections
016.20	TB OF URETER-UNSPEC	2		Other infections
016.21	TB OF URETER-NO EXAM	2		Other infections
016.22	TB OF URETER-EXAM UNKN	2		Other infections
016.23	TB OF URETER-MICRO DX	2		Other infections
016.24	TB OF URETER-CULT DX	2		Other infections
016.25	TB OF URETER-HISTO DX	2		Other infections
016.26	TB OF URETER-OTH TEST	2		Other infections
016.3	TB OF URINARY ORGAN NEC*	2		Other infections
016.30	TB URINARY NEC-UNSPEC	2		Other infections
016.31	TB URINARY NEC-NO EXAM	2		Other infections
016.32	TB URINARY NEC-EXAM UNKN	2		Other infections

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
016.33	TB URINARY NEC-MICRO DX	2		Other infections
016.34	TB URINARY NEC-CULT DX	2		Other infections
016.35	TB URINARY NEC-HISTO DX	2		Other infections
016.36	TB URINARY NEC-OTH TEST	2		Other infections
016.4	TB OF EPIDIDYMIS*	2		Other infections
016.40	TB EPIDIDYMIS-UNSPEC	2		Other infections
016.41	TB EPIDIDYMIS-NO EXAM	2		Other infections
016.42	TB EPIDIDYMIS-EXAM UNKN	2		Other infections
016.43	TB EPIDIDYMIS-MICRO DX	2		Other infections
016.44	TB EPIDIDYMIS-CULT DX	2		Other infections
016.45	TB EPIDIDYMIS-HISTO DX	2		Other infections
016.46	TB EPIDIDYMIS-OTH TEST	2		Other infections
016.5	TB MALE GENITAL ORG NEC*	2		Other infections
016.50	TB MALE GENIT NEC-UNSPEC	2		Other infections
016.51	TB MALE GEN NEC-NO EXAM	2		Other infections
016.52	TB MALE GEN NEC-EX UNKN	2		Other infections
016.53	TB MALE GEN NEC-MICRO DX	2		Other infections
016.54	TB MALE GEN NEC-CULT DX	2		Other infections
016.55	TB MALE GEN NEC-HISTO DX	2		Other infections
016.56	TB MALE GEN NEC-OTH TEST	2		Other infections
016.6	TB OF OVARY AND TUBE*	2		Other infections
016.60	TB OVARY & TUBE-UNSPEC	2		Other infections
016.61	TB OVARY & TUBE-NO EXAM	2		Other infections
016.62	TB OVARY/TUBE-EXAM UNKN	2		Other infections
016.63	TB OVARY & TUBE-MICRO DX	2		Other infections
016.64	TB OVARY & TUBE-CULT DX	2		Other infections
016.65	TB OVARY & TUBE-HISTO DX	2		Other infections
016.66	TB OVARY & TUBE-OTH TEST	2		Other infections
016.7	TB FEMALE GENIT ORG NEC*	2		Other infections
016.70	TB FEMALE GEN NEC-UNSPEC	2		Other infections
016.71	TB FEM GEN NEC-NO EXAM	2		Other infections
016.72	TB FEM GEN NEC-EXAM UNKN	2		Other infections
016.73	TB FEM GEN NEC-MICRO DX	2		Other infections
016.74	TB FEM GEN NEC-CULT DX	2		Other infections
016.75	TB FEM GEN NEC-HISTO DX	2		Other infections
016.76	TB FEM GEN NEC-OTH TEST	2		Other infections
016.9	GENITOURINARY TB NOS*	2		Other infections
016.90	GU TB NOS-UNSPEC	2		Other infections
016.91	GU TB NOS-NO EXAM	2		Other infections
016.92	GU TB NOS-EXAM UNKN	2		Other infections
016.93	GU TB NOS-MICRO DX	2		Other infections
016.94	GU TB NOS-CULT DX	2		Other infections
016.95	GU TB NOS-HISTO DX	2		Other infections
016.96	GU TB NOS-OTH TEST	2		Other infections
017.	TUBERCULOSIS NEC*	2		Other infections

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
017.0	TB SKIN & SUBCUTANEOUS*	2		Other infections
017.00	TB SKIN/SUBCUTAN-UNSPEC	2		Other infections
017.01	TB SKIN/SUBCUT-NO EXAM	2		Other infections
017.02	TB SKIN/SUBCUT-EXAM UNKN	2		Other infections
017.03	TB SKIN/SUBCUT-MICRO DX	2		Other infections
017.04	TB SKIN/SUBCUT-CULT DX	2		Other infections
017.05	TB SKIN/SUBCUT-HISTO DX	2		Other infections
017.06	TB SKIN/SUBCUT-OTH TEST	2		Other infections
017.1	ERYTHEMA NODOSUM IN TB*	2		Other infections
017.10	ERYTHEMA NODOS TB-UNSPEC	2		Other infections
017.11	ERYTHEM NODOS TB-NO EXAM	2		Other infections
017.12	ERYTHEM NOD TB-EXAM UNKN	2		Other infections
017.13	ERYTHEM NOD TB-MICRO DX	2		Other infections
017.14	ERYTHEM NODOS TB-CULT DX	2		Other infections
017.15	ERYTHEM NOD TB-HISTO DX	2		Other infections
017.16	ERYTHEM NOD TB-OTH TEST	2		Other infections
017.2	TB OF PERIPH LYMPH NODE*	2		Other infections
017.20	TB PERIPH LYMPH-UNSPEC	2		Other infections
017.21	TB PERIPH LYMPH-NO EXAM	2		Other infections
017.22	TB PERIPH LYMPH-EXAM UNK	2		Other infections
017.23	TB PERIPH LYMPH-MICRO DX	2		Other infections
017.24	TB PERIPH LYMPH-CULT DX	2		Other infections
017.25	TB PERIPH LYMPH-HISTO DX	2		Other infections
017.26	TB PERIPH LYMPH-OTH TEST	2		Other infections
017.3	TB OF EYE*	2		Other infections
017.30	TB OF EYE-UNSPEC	2		Other infections
017.31	TB OF EYE-NO EXAM	2		Other infections
017.32	TB OF EYE-EXAM UNKN	2		Other infections
017.33	TB OF EYE-MICRO DX	2		Other infections
017.34	TB OF EYE-CULT DX	2		Other infections
017.35	TB OF EYE-HISTO DX	2		Other infections
017.36	TB OF EYE-OTH TEST	2		Other infections
017.4	TB OF EAR*	2		Other infections
017.40	TB OF EAR-UNSPEC	2		Other infections
017.41	TB OF EAR-NO EXAM	2		Other infections
017.42	TB OF EAR-EXAM UNKN	2		Other infections
017.43	TB OF EAR-MICRO DX	2		Other infections
017.44	TB OF EAR-CULT DX	2		Other infections
017.45	TB OF EAR-HISTO DX	2		Other infections
017.46	TB OF EAR-OTH TEST	2		Other infections
017.5	TB OF THYROID GLAND*	2		Other infections
017.50	TB OF THYROID-UNSPEC	2		Other infections
017.51	TB OF THYROID-NO EXAM	2		Other infections
017.52	TB OF THYROID-EXAM UNKN	2		Other infections
017.53	TB OF THYROID-MICRO DX	2		Other infections

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
017.54	TB OF THYROID-CULT DX	2		Other infections
017.55	TB OF THYROID-HISTO DX	2		Other infections
017.56	TB OF THYROID-OTH TEST	2		Other infections
017.6	TB OF ADRENAL GLAND*	2		Other infections
017.60	TB OF ADRENAL-UNSPEC	2		Other infections
017.61	TB OF ADRENAL-NO EXAM	2		Other infections
017.62	TB OF ADRENAL-EXAM UNKN	2		Other infections
017.63	TB OF ADRENAL-MICRO DX	2		Other infections
017.64	TB OF ADRENAL-CULT DX	2		Other infections
017.65	TB OF ADRENAL-HISTO DX	2		Other infections
017.7	TB OF SPLEEN*	2		Other infections
017.70	TB OF SPLEEN-UNSPEC	2		Other infections
017.71	TB OF SPLEEN-NO EXAM	2		Other infections
017.72	TB OF SPLEEN-EXAM UNKN	2		Other infections
017.73	TB OF SPLEEN-MICRO DX	2		Other infections
017.74	TB OF SPLEEN-CULT DX	2		Other infections
017.75	TB OF SPLEEN-HISTO DX	2		Other infections
017.76	TB OF SPLEEN-OTH TEST	2		Other infections
017.8	TB OF ESOPHAGUS*	2		Other infections
017.80	TB ESOPHAGUS-UNSPEC	2		Other infections
017.81	TB ESOPHAGUS-NO EXAM	2		Other infections
017.82	TB ESOPHAGUS-EXAM UNKN	2		Other infections
017.83	TB ESOPHAGUS-MICRO DX	2		Other infections
017.84	TB ESOPHAGUS-CULT DX	2		Other infections
017.85	TB ESOPHAGUS-HISTO DX	2		Other infections
017.86	TB ESOPHAGUS-OTH TEST	2		Other infections
017.9	TB OF ORGAN NEC*	2		Other infections
017.90	TB OF ORGAN NEC-UNSPEC	2		Other infections
017.91	TB OF ORGAN NEC-NO EXAM	2		Other infections
017.92	TB ORGAN NEC-EXAM UNKN	2		Other infections
017.93	TB OF ORGAN NEC-MICRO DX	2		Other infections
017.94	TB OF ORGAN NEC-CULT DX	2		Other infections
017.95	TB OF ORGAN NEC-HISTO DX	2		Other infections
017.96	TB OF ORGAN NEC-OTH TEST	2		Other infections
018.	MILIARY TUBERCULOSIS*	2		Other infections
018.0	ACUTE MILIARY TB*	2		Other infections
018.00	ACUTE MILIARY TB-UNSPEC	2		Other infections
018.01	ACUTE MILIARY TB-NO EXAM	2		Other infections
018.02	AC MILIARY TB-EXAM UNKN	2		Other infections
018.03	AC MILIARY TB-MICRO DX	2		Other infections
018.04	ACUTE MILIARY TB-CULT DX	2		Other infections
018.05	AC MILIARY TB-HISTO DX	2		Other infections
018.06	AC MILIARY TB-OTH TEST	2		Other infections
018.8	MILIARY TB NEC*	2		Other infections
018.80	MILIARY TB NEC-UNSPEC	2		Other infections

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
018.81	MILIARY TB NEC-NO EXAM	2		Other infections
018.82	MILIARY TB NEC-EXAM UNKN	2		Other infections
018.83	MILIARY TB NEC-MICRO DX	2		Other infections
018.84	MILIARY TB NEC-CULT DX	2		Other infections
018.85	MILIARY TB NEC-HISTO DX	2		Other infections
018.86	MILIARY TB NEC-OTH TEST	2		Other infections
018.9	MILIARY TUBERCULOSIS NOS*	2		Other infections
018.90	MILIARY TB NOS-UNSPEC	2		Other infections
018.91	MILIARY TB NOS-NO EXAM	2		Other infections
018.92	MILIARY TB NOS-EXAM UNKN	2		Other infections
018.93	MILIARY TB NOS-MICRO DX	2		Other infections
018.94	MILIARY TB NOS-CULT DX	2		Other infections
018.95	MILIARY TB NOS-HISTO DX	2		Other infections
018.96	MILIARY TB NOS-OTH TEST	2		Other infections
027.0	LISTERIOSIS	2		Other infections
027.1	ERYSIPELOTHRIX INFECTION	2		Other infections
027.2	PASTEURELLOSIS	2		Other infections
027.8	ZOONOTIC BACT DIS NEC	2		Other infections
027.9	ZOONOTIC BACT DIS NOS	2		Other infections
036.0	MENINGOCOCCAL MENINGITIS	2	3,5	Meningitis and encephalitis
036.2	MENINGOCOCCEMIA	3	3,5	"Major" comorbidities
036.3	MENINGOCOCC ADRENAL SYND	3	5	"Major" comorbidities
036.40	MENINGOCOCC CARDITIS NOS	3	14	"Major" comorbidities
036.42	MENINGOCOCC ENDOCARDITIS	3	14	"Major" comorbidities
036.43	MENINGOCOCC MYOCARDITIS	3	14	"Major" comorbidities
037.	TETANUS	3	6	"Major" comorbidities
038.0	STREPTOCOCCAL SEPTICEMIA	2		Other infections
038.1	STAPHYLOCOCC SEPTICEMIA*	2		Other infections
038.10	STAPHYLCOCC SEPTICEM NOS	2		Other infections
038.11	STAPH AUREUS SEPTICEMIA	2		Other infections
038.19	STAPHYLCOCC SEPTICEM NEC	2		Other infections
038.2	PNEUMOCOCCAL SEPTICEMIA	2		Other infections
038.3	ANAEROBIC SEPTICEMIA	2		Other infections
038.4	GRAM-NEG SEPTICEMIA NEC*	2		Other infections
038.40	GRAM-NEG SEPTICEMIA NOS	2		Other infections
038.41	H. INFLUENAE SEPTICEMIA	2		Other infections
038.42	E COLI SEPTICEMIA	2		Other infections
038.43	PSEUDOMONAS SEPTICEMIA	2		Other infections
038.44	SERRATIA SEPTICEMIA	2		Other infections
038.49	GRAM-NEG SEPTICEMIA NEC	2		Other infections
038.8	SEPTICEMIA NEC	2		Other infections
038.9	SEPTICEMIA NOS	2		Other infections
041.7	PSEUDOMONAS INFECT NOS	2		Pseudomonas
042.	HUMAN IMMUNO VIRUS DIS	2		Other infections
047.8	VIRAL MENINGITIS NEC	2	3,5	Meningitis and encephalitis

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
047.9	VIRAL MENINGITIS NOS	2	3,5	Meningitis and encephalitis
048.	OTH ENTEROVIRAL CNS DIS	2	3,5	Meningitis and encephalitis
049.0	LYMPHOCYTIC CHORIOMENING	2	3,5	Meningitis and encephalitis
049.9	VIRAL ENCEPHALITIS NOS	2	3	Meningitis and encephalitis
052.0	POSTVARICELLA ENCEPHALIT	2	3	Meningitis and encephalitis
052.1	VARICELLA PNEUMONITIS	3	15	"Major" comorbidities
053.0	HERPES ZOSTER MENINGITIS	2	3,5	Meningitis and encephalitis
053.13	POSTHERPES POLYNEUROPATH	2	6	Meningitis and encephalitis
054.3	HERPETIC ENCEPHALITIS	2	3	Meningitis and encephalitis
054.5	HERPETIC SEPTICEMIA	2	3	Other infections
054.72	H SIMPLEX MENINGITIS	2	3,5	Meningitis and encephalitis
054.79	H SIMPLEX COMPLICAT NEC	3		"Major" comorbidities
055.0	POSTMEASLES ENCEPHALITIS	2	3	Meningitis and encephalitis
055.1	POSTMEASLES PNEUMONIA	3	15	"Major" comorbidities
070.20	HPT B ACTE COMA WO DLTA	3	3	"Major" comorbidities
070.21	HPT B ACTE COMA W DLTA	3	3	"Major" comorbidities
070.22	HPT B CHRN COMA WO DLTA	3	3	"Major" comorbidities
070.23	HPT B CHRN COMA W DLTA	3	3	"Major" comorbidities
070.41	HPT C ACUTE W HEPAT COMA	3	3	"Major" comorbidities
070.42	HPT DLT WO B W HPT COMA	3	3	"Major" comorbidities
070.43	HPT E W HEPAT COMA	3	3	"Major" comorbidities
070.44	CHRNC HPT C W HEPAT COMA	3	3	"Major" comorbidities
070.49	OTH VRL HEPAT W HPT COMA	3	3	"Major" comorbidities
070.6	VIRAL HEPAT NOS W COMA	3	3	"Major" comorbidities
072.1	MUMPS MENINGITIS	2	3,5	Meningitis and encephalitis
072.2	MUMPS ENCEPHALITIS	2	3	Meningitis and encephalitis
072.3	MUMPS PANCREATITIS	3		"Major" comorbidities
079.50	RETROVIRUS_ UNSPECIFIED	2		Other infections
079.51	HTLV-1 INFECTION OTH DIS	2	6	Other infections
079.52	HTLV-II INFECTN OTH DIS	2	6	Other infections
079.53	HIV-2 INFECTION OTH DIS	2		Other infections
079.59	OTH SPECFIED RETROVIRUS	2		Other infections
090.42	CONGEN SYPH MENINGITIS	2	3,5	Meningitis and encephalitis
093.20	SYPHIL ENDOCARDITIS NOS	3	14	"Major" comorbidities
093.82	SYPHILITIC MYOCARDITIS	3	14	"Major" comorbidities
094.2	SYPHILITIC MENINGITIS	2	3,5	Meningitis and encephalitis
094.87	SYPH RUPT CEREB ANEURYSM	3	1,3	"Major" comorbidities
098.89	GONOCOCCAL INF SITE NEC	2		Other infections
112.4	CANDIDIASIS OF LUNG	1	15	Candidiasis (selected)
112.5	DISSEMINATED CANDIDIASIS	1		Candidiasis (selected)
112.81	CANDIDAL ENDOCARDITIS	1	14	Candidiasis (selected)
112.83	CANDIDAL MENINGITIS	1	3,5	Candidiasis (selected)
112.84	CANDIDAL ESOPHAGITIS	1		Candidiasis (selected)
114.2	COCCIDIOIDAL MENINGITIS	2	3,5	Meningitis and encephalitis
115.	HISTOPLASMOSIS*	2	15	Other infections

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
115.0	HISTOPLASMA CAPSULATUM*	2	15	Other infections
115.00	HISTOPLASMA CAPSULAT NOS	2	15	Other infections
115.01	HISTOPLASM CAPSUL MENING	2	3,5	Other infections
115.02	HISTOPLASM CAPSUL RETINA	2		Other infections
115.03	HISTOPLASM CAPS PERICARD	2	14	Other infections
115.04	HISTOPLASM CAPS ENDOCARD	2	14	Other infections
115.05	HISTOPLASM CAPS PNEUMON	2	15	Other infections
115.09	HISTOPLASMA CAPSULAT NEC	2	15	Other infections
115.1	HISTOPLASMA DUBOISII*	2	15	Other infections
115.10	HISTOPLASMA DUBOISII NOS	2		Other infections
115.11	HISTOPLASM DUBOIS MENING	2	3,5	Other infections
115.12	HISTOPLASM DUBOIS RETINA	2		Other infections
115.13	HISTOPLASM DUB PERICARD	2	14	Other infections
115.14	HISTOPLASM DUB ENDOCARD	2	14	Other infections
115.15	HISTOPLASM DUB PNEUMONIA	2	15	Other infections
115.19	HISTOPLASMA DUBOISII NEC	2	15	Other infections
115.9	HISTOPLASMOSIS_ UNSPEC*	2	15	Other infections
115.90	HISTOPLASMOSIS NOS	2	15	Other infections
115.91	HISTOPLASMOSIS MENINGIT	2	3,5	Other infections
115.92	HISTOPLASMOSIS RETINITIS	2		Other infections
115.93	HISTOPLASMOSIS PERICARD	2	14	Other infections
115.94	HISTOPLASMOSIS ENDOCARD	2	14	Other infections
115.95	HISTOPLASMOSIS PNEUMONIA	2	15	Other infections
115.99	HISTOPLASMOSIS NEC	2	15	Other infections
130.0	TOXOPLASM MENINGOENCEPH	2	3,5	Meningitis and encephalitis
130.3	TOXOPLASMA MYOCARDITIS	3	14	"Major" comorbidities
130.4	TOXOPLASMA PNEUMONITIS	3	15	"Major" comorbidities
136.3	PNEUMOCYSTOSIS	3	15	"Major" comorbidities
139.0	LATE EFF VIRAL ENCEPHAL	2	3	Meningitis and encephalitis
204.00	ACT LYM LEUK W/O RMSION	3		"Major" comorbidities
205.00	ACT MYL LEUK W/O RMSION	3		"Major" comorbidities
206.00	ACT MONO LEUK W/O RMSION	3		"Major" comorbidities
207.00	ACT ERTH/ERYLK W/O RMSON	3		"Major" comorbidities
208.00	ACT LEUK UNS CL W/O RMSN	3		"Major" comorbidities
235.1	UNC BEHAV NEO ORAL/PHAR	1		Misc. throat problems
250.40	DMII RENL NT ST UNCNTRLD	3		Renal complications of diabetes
250.41	DMI RENL NT ST UNCNTRLD	3		Renal complications of diabetes
250.42	DMII RENAL UNCNTRLD	3		Renal complications of diabetes
250.43	DMI RENAL UNCNTRLD	3		Renal complications of diabetes
250.50	DMII OPHTH NT ST UNCNTRL	3		Non-renal complications of diabetes
250.51	DMI OPHTH NT ST UNCNTRLD	3		Non-renal complications of diabetes

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
250.52	DMII OPHTH UNCNTRLD	3		Non-renal complications of diabetes
250.53	DMI OPHTH UNCNTRLD	3		Non-renal complications of diabetes
250.60	DMII NEURO NT ST UNCNTRL	3	6	Non-renal complications of diabetes
250.61	DMI NEURO NT ST UNCNTRLD	3	6	Non-renal complications of diabetes
250.62	DMII NEURO UNCNTRLD	3	6	Non-renal complications of diabetes
250.63	DMI NEURO UNCNTRLD	3	6	Non-renal complications of diabetes
250.70	DMII CIRC NT ST UNCNTRLD	3		Non-renal complications of diabetes
250.71	DMI CIRC NT ST UNCNTRLD	3		Non-renal complications of diabetes
250.72	DMII CIRC UNCNTRLD	3		Non-renal complications of diabetes
250.73	DMI CIRC UNCNTRLD	3		Non-renal complications of diabetes
250.80	DMII OTH NT ST UNCNTRLD	3		Non-renal complications of diabetes
250.81	DMI OTH NT ST UNCNTRLD	3		Non-renal complications of diabetes
250.82	DMII OTH UNCNTRLD	3		Non-renal complications of diabetes
250.83	DMI OTH UNCNTRLD	3		Non-renal complications of diabetes
250.90	DMII UNSPF NT ST UNCNTRL	3		Non-renal complications of diabetes
250.91	DMI UNSPF NT ST UNCNTRLD	3		Non-renal complications of diabetes
250.92	DMII UNSPF UNCNTRLD	3		Non-renal complications of diabetes
250.93	DMI UNSPF UNCNTRLD	3		Non-renal complications of diabetes
260.	KWASHIORKOR	1		Malnutrition
261.	NUTRITIONAL MARASMUS	1		Malnutrition
262.	OTH SEVERE MALNUTRITION	1		Malnutrition
277.00	CYSTIC FIBROS W/O ILEUS	3	15	"Major" comorbidities
277.01	CYSTIC FIBROSIS W ILEUS	3	15	"Major" comorbidities
278.01	MORBID OBESITY	3		Obesity
282.60	SICKLE-CELL ANEMIA NOS	3		Aplastic anemia and selected anemias
282.61	HB-S DISEASE W/O CRISIS	3		Aplastic anemia and selected anemias
282.62	HB-S DISEASE WITH CRISIS	3		Aplastic anemia and selected anemias
282.63	SICKLE-CELL/HB-C DISEASE	3		Aplastic anemia and selected anemias

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
282.69	SICKLE-CELL ANEMIA NEC	3		Aplastic anemia and selected anemias
284.0	CONGEN APLASTIC ANEMIA	3		Aplastic anemia and selected anemias
284.8	APLASTIC ANEMIAS NEC	3		Aplastic anemia and selected anemias
284.9	APLASTIC ANEMIA NOS	3		Aplastic anemia and selected anemias
286.0	CONG FACTOR VIII DIORD	3		"Major" comorbidities
286.1	CONG FACTOR IX DISORDER	3		"Major" comorbidities
286.6	DEFIBRINATION SYNDROME	3		"Major" comorbidities
320.0	HEMOPHILUS MENINGITIS	2	3,5	Meningitis and encephalitis
320.1	PNEUMOCOCCAL MENINGITIS	2	3,5	Meningitis and encephalitis
320.2	STREPTOCOCCAL MENINGITIS	2	3,5	Meningitis and encephalitis
320.3	STAPHYLOCOCC MENINGITIS	2	3,5	Meningitis and encephalitis
320.7	MENING IN OTH BACT DIS	2	3,5	Meningitis and encephalitis
320.81	ANAEROBIC MENINGITIS	2	3,5	Meningitis and encephalitis
320.82	MNINGTS GRAM-NEG BCT NEC	2	3,5	Meningitis and encephalitis
320.89	MENINGITIS OTH SPCF BACT	2	3,5	Meningitis and encephalitis
320.9	BACTERIAL MENINGITIS NOS	2	3,5	Meningitis and encephalitis
321.0	CRYPTOCOCCAL MENINGITIS	2	3,5	Meningitis and encephalitis
321.1	MENING IN OTH FUNGAL DIS	2	3,5	Meningitis and encephalitis
321.2	MENING IN OTH VIRAL DIS	2	3,5	Meningitis and encephalitis
321.3	TRYPANOSOMIASIS MENINGIT	2	3,5	Meningitis and encephalitis
321.4	MENINGIT D/T SARCOIDOSIS	2	3,5	Meningitis and encephalitis
321.8	MENING IN OTH NONBAC DIS	2	3,5	Meningitis and encephalitis
322.0	NONPYOGENIC MENINGITIS	2	3,5	Meningitis and encephalitis
322.2	CHRONIC MENINGITIS	2	3,5	Meningitis and encephalitis
322.9	MENINGITIS NOS	2	3,5	Meningitis and encephalitis
323.6	POSTINFECT ENCEPHALITIS	2	3	Meningitis and encephalitis
323.8	ENCEPHALITIS NEC	2	3	Meningitis and encephalitis
323.9	ENCEPHALITIS NOS	2	3	Meningitis and encephalitis
324.0	INTRACRANIAL ABSCESS	3	3	"Major" comorbidities
324.1	INTRASPINAL ABSCESS	3	3	"Major" comorbidities
324.9	CNS ABSCESS NOS	3	3	"Major" comorbidities
342.00	FLCCD HMIPLGA UNSPF SIDE	3	1	Hemiplegia
342.01	FLCCD HMIPLGA DOMNT SIDE	3	1	Hemiplegia
342.02	FLCCD HMIPLG NONDMNT SDE	3	1	Hemiplegia
342.10	SPSTC HMIPLGA UNSPF SIDE	3	1	Hemiplegia
342.11	SPSTC HMIPLGA DOMNT SIDE	3	1	Hemiplegia
342.12	SPSTC HMIPLG NONDMNT SDE	3	1	Hemiplegia
342.80	OT SP HMIPLGA UNSPF SIDE	3	1	Hemiplegia
342.81	OT SP HMIPLGA DOMNT SIDE	3	1	Hemiplegia
342.82	OT SP HMIPLG NONDMNT SDE	3	1	Hemiplegia
342.90	UNSP HEMIPLGA UNSPF SIDE	3	1	Hemiplegia

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
342.91	UNSP HEMIPLGA DOMNT SIDE	3	1	Hemiplegia
342.92	UNSP HMIPLGA NONDMNT SDE	3	1	Hemiplegia
345.11	GEN CNV EPIL W INTR EPIL	3	2,3	"Major" comorbidities
345.3	GRAND MAL STATUS	3	2,3	"Major" comorbidities
348.1	ANOXIC BRAIN DAMAGE	3	2,3	"Major" comorbidities
356.4	IDIO PROG POLYNEUROPATHY	2	3,6,19	Meningitis and encephalitis
357.2	NEUROPATHY IN DIABETES	3	6	Non-renal complications of diabetes
376.01	ORBITAL CELLULITIS	2		Other infections
376.02	ORBITAL PERIOSTITIS	3		"Major" comorbidities
376.03	ORBITAL OSTEOMYELITIS	3		"Major" comorbidities
398.0	RHEUMATIC MYOCARDITIS	3	14	"Major" comorbidities
403.01	MAL HYP REN W RENAL FAIL	3		"Major" comorbidities
404.01	MAL HYPER HRT/REN W CHF	3	14	"Major" comorbidities
404.03	MAL HYP HRT/REN W CHF&RF	3	14	"Major" comorbidities
410.01	AMI ANTEROLATERAL_ INIT	3	14	"Major" comorbidities
410.11	AMI ANTERIOR WALL_ INIT	3	14	"Major" comorbidities
410.21	AMI INFEROLATERAL_ INIT	3	14	"Major" comorbidities
410.31	AMI INFEROPOST_ INITIAL	3	14	"Major" comorbidities
410.41	AMI INFERIOR WALL_ INIT	3	14	"Major" comorbidities
410.51	AMI LATERAL NEC_ INITIAL	3	14	"Major" comorbidities
410.61	TRUE POST INFARCT_ INIT	3	14	"Major" comorbidities
410.71	SUBENDO INFARCT_ INITIAL	3	14	"Major" comorbidities
410.81	AMI NEC_ INITIAL	3	14	"Major" comorbidities
410.91	AMI NOS_ INITIAL	3	14	"Major" comorbidities
415.1	PULMON EMBOLISM/INFARCT*	3	15	"Major" comorbidities
415.11	IATROGEN PULM EMB/INFARC	3	15	"Major" comorbidities
415.19	PULM EMBOL/INFARCT NEC	3	15	"Major" comorbidities
421.0	AC/SUBAC BACT ENDOCARD	3	14	"Major" comorbidities
421.1	AC ENDOCARDIT IN OTH DIS	3	14	"Major" comorbidities
421.9	AC/SUBAC ENDOCARDIT NOS	3	14	"Major" comorbidities
422.0	AC MYOCARDIT IN OTH DIS	3	14	"Major" comorbidities
422.90	ACUTE MYOCARDITIS NOS	3	14	"Major" comorbidities
422.91	IDIOPATHIC MYOCARDITIS	3	14	"Major" comorbidities
422.92	SEPTIC MYOCARDITIS	3	14	"Major" comorbidities
422.93	TOXIC MYOCARDITIS	3	14	"Major" comorbidities
422.99	ACUTE MYOCARDITIS NEC	3	14	"Major" comorbidities
427.41	VENTRICULAR FIBRILLATION	3	14	"Major" comorbidities
427.5	CARDIAC ARREST	3	14	"Major" comorbidities
430.	SUBARACHNOID HEMORRHAGE	3	1,2,3	"Major" comorbidities
431.	INTRACEREBRAL HEMORRHAGE	3	1,2,3	"Major" comorbidities
432.0	NONTRAUM EXTRADURAL HEM	3	1,2,3	"Major" comorbidities
432.1	SUBDURAL HEMORRHAGE	3	1,2,3	"Major" comorbidities
433.01	OCL BSLR ART W INFRCT	3	1	"Major" comorbidities
433.11	OCL CRTD ART W INFRCT	3	1	"Major" comorbidities

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
433.21	OCL VRTB ART W INFRCT	3	1	"Major" comorbidities
433.31	OCL MLT BI ART W INFRCT	3	1	"Major" comorbidities
433.81	OCL SPCF ART W INFRCT	3	1	"Major" comorbidities
433.91	OCL ART NOS W INFRCT	3	1	"Major" comorbidities
434.01	CRBL THRMBS W INFRCT	3	1	"Major" comorbidities
434.11	CRBL EMBLSM W INFRCT	3	1	"Major" comorbidities
434.91	CRBL ART OCL NOS W INFRC	3	1	"Major" comorbidities
436.	CVA	3	1	"Major" comorbidities
438.82	LATE EF CV DIS DYSPHAGIA	2	1	Dysphagia
440.23	ATH EXT NTV ART ULCRTION	3	10,11	"Major" comorbidities
440.24	ATH EXT NTV ART GNGRENE	3	10,11	"Major" comorbidities
441.0	DISSECTING ANEURYSM*	3		"Major" comorbidities
441.00	DSCT OF AORTA UNSP SITE	3		"Major" comorbidities
441.01	DSCT OF THORACIC AORTA	3	5	"Major" comorbidities
441.02	DSCT OF ABDOMINAL AORTA	3	5	"Major" comorbidities
441.03	DSCT OF THORACOABD AORTA	3	5	"Major" comorbidities
441.1	RUPTUR THORACIC ANEURYSM	3	5	"Major" comorbidities
441.3	RUPT ABD AORTIC ANEURYSM	3	5	"Major" comorbidities
441.5	RUPT AORTIC ANEURYSM NOS	3	5	"Major" comorbidities
441.6	THORACOABD ANEURYSM RUPT	3	5	"Major" comorbidities
446.3	LETHAL MIDLINE GRANULOMA	3		"Major" comorbidities
452.	PORTAL VEIN THROMBOSIS	3		"Major" comorbidities
453.	OTH VENOUS THROMBOSIS*	3		"Major" comorbidities
453.0	BUDD-CHIARI SYNDROME	3		"Major" comorbidities
453.1	THROMBOPHLEBITIS MIGRANS	3		"Major" comorbidities
453.2	VENA CAVA THROMBOSIS	3		"Major" comorbidities
453.3	RENAL VEIN THROMBOSIS	3		"Major" comorbidities
464.11	AC TRACHEITIS W OBSTRUCT	3	15	"Major" comorbidities
464.21	AC LARYNGOTRACH W OBSTR	3	15	"Major" comorbidities
464.31	AC EPIGLOTTITIS W OBSTR	3	15	"Major" comorbidities
466.1	ACUTE BRONCHIOLITIS*	3	15	"Major" comorbidities
478.30	VOCAL CORD PARALYSIS NOS	1	15	Vocal cord paralysis
478.31	VOCAL PARAL UNILAT PART	1	15	Vocal cord paralysis
478.32	VOCAL PARAL UNILAT TOTAL	1	15	Vocal cord paralysis
478.33	VOCAL PARAL BILAT PART	1	15	Vocal cord paralysis
478.34	VOCAL PARAL BILAT TOTAL	1	15	Vocal cord paralysis
478.6	EDEMA OF LARYNX	1	15	Vocal cord paralysis
480.0	ADENOVIRAL PNEUMONIA	3	15	Pneumonia
480.1	RESP SYNCYT VIRAL PNEUM	3	15	Pneumonia
480.2	PARINFLUENZA VIRAL PNEUM	3	15	Pneumonia
480.8	VIRAL PNEUMONIA NEC	3	15	Pneumonia
480.9	VIRAL PNEUMONIA NOS	3	15	Pneumonia
481.	PNEUMOCOCCAL PNEUMONIA	3	15	Pneumonia
482.0	K. PNEUMONIAE PNEUMONIA	3	15	Pneumonia
482.1	PSEUDOMONAL PNEUMONIA	3	15	Pneumonia

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
482.2	H.INFLUENZAE PNEUMONIA	3	15	Pneumonia
482.30	STREPTOCOCCAL PNEUMN NOS	3	15	Pneumonia
482.31	PNEUMONIA STRPTOCOCCUS A	3	15	Pneumonia
482.32	PNEUMONIA STRPTOCOCCUS B	3	15	Pneumonia
482.39	PNEUMONIA OTH STREP	3	15	Pneumonia
482.40	STAPHYLOCOCCAL PNEU NOS	3	15	Pneumonia
482.41	STAPH AUREUS PNEUMONIA	3	15	Pneumonia
482.49	STAPH PNEUMONIA NEC	3	15	Pneumonia
482.8	BACTERIAL PNEUMONIA NEC*	3	15	Pneumonia
482.81	PNEUMONIA ANAEROBES	3	15	Pneumonia
482.82	PNEUMONIA E COLI	3	15	Pneumonia
482.83	PNEUMO OTH GRM-NEG BACT	3	15	Pneumonia
482.84	LEGIONNAIRES' DISEASE	3	15	Pneumonia
482.89	PNEUMONIA OTH SPCF BACT	3	15	Pneumonia
482.9	BACTERIAL PNEUMONIA NOS	3	15	Pneumonia
483.0	PNEU MYCPLSM PNEUMONIAE	3	15	Pneumonia
483.1	PNEUMONIA D/T CHLAMYDIA	3	15	Pneumonia
483.8	PNEUMON OTH SPEC ORGNSM	3	15	Pneumonia
484.1	PNEUM W CYTOMEG INCL DIS	3	15	Pneumonia
484.3	PNEUMONIA IN WHOOP COUGH	3	15	Pneumonia
484.5	PNEUMONIA IN ANTHRAX	3	15	Pneumonia
484.6	PNEUM IN ASPERGILLOSIS	3	15	Pneumonia
484.7	PNEUM IN OTH SYS MYCOSES	3	15	Pneumonia
484.8	PNEUM IN INFECT DIS NEC	3	15	Pneumonia
485.	BRONCHOPNEUMONIA ORG NOS	3	15	Pneumonia
486.	PNEUMONIA_ ORGANISM NOS	3	15	Pneumonia
487.0	INFLUENZA WITH PNEUMONIA	3	15	Pneumonia
506.0	FUM/VAPOR BRONC/PNEUMON	3	15	"Major" comorbidities
506.1	FUM/VAPOR AC PULM EDEMA	3	15	"Major" comorbidities
507.0	FOOD/VOMIT PNEUMONITIS	3	15	"Major" comorbidities
507.1	OIL/ESSENCE PNEUMONITIS	3	15	"Major" comorbidities
507.8	SOLID/LIQ PNEUMONIT NEC	3	15	"Major" comorbidities
510.0	EMPYEMA WITH FISTULA	3	15	"Major" comorbidities
510.9	EMPYEMA W/O FISTULA	3	15	"Major" comorbidities
511.1	BACT PLEUR/EFFUS NOT TB	3	15	"Major" comorbidities
513.0	ABSCESS OF LUNG	3	15	"Major" comorbidities
513.1	ABSCESS OF MEDIASTINUM	3	15	"Major" comorbidities
514.	PULM CONGEST/HYPOSTASIS	3	15	"Major" comorbidities
515.	POSTINFLAM PULM FIBROSIS	3	15	"Major" comorbidities
518.3	PULMONARY EOSINOPHILIA	3	15	"Major" comorbidities
518.5	POST TRAUM PULM INSUFFIC	3	15	"Major" comorbidities
518.81	ACUTE RESPIRATRY FAILURE	3	15	"Major" comorbidities
519.2	MEDIASTINITIS	3	15	"Major" comorbidities
528.3	CELLULITIS/ABSCESS MOUTH	2		Other infections
530.0	ACHALASIA & CARDIOSPASM	3		Esophageal conditions

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
530.3	ESOPHAGEAL STRICTURE	3		Esophageal conditions
530.4	PERFORATION OF ESOPHAGUS	3	15	"Major" comorbidities
530.6	ACQ ESOPHAG DIVERTICULUM	3		Esophagal conditions
530.82	ESOPHAGEAL HEMORRHAGE	3		"Major" comorbidities
531.00	AC STOMACH ULCER W HEM	3		"Major" comorbidities
531.01	AC STOMAC ULC W HEM-OBST	3		"Major" comorbidities
531.10	AC STOMACH ULCER W PERF	3		"Major" comorbidities
531.11	AC STOM ULC W PERF-OBST	3		"Major" comorbidities
531.20	AC STOMAC ULC W HEM/PERF	3		"Major" comorbidities
531.21	AC STOM ULC HEM/PERF-OBS	3		"Major" comorbidities
531.40	CHR STOMACH ULC W HEM	3		"Major" comorbidities
531.41	CHR STOM ULC W HEM-OBSTR	3		"Major" comorbidities
531.50	CHR STOMACH ULCER W PERF	3		"Major" comorbidities
531.51	CHR STOM ULC W PERF-OBST	3		"Major" comorbidities
531.60	CHR STOMACH ULC HEM/PERF	3		"Major" comorbidities
531.61	CHR STOM ULC HEM/PERF-OB	3		"Major" comorbidities
532.00	AC DUODENAL ULCER W HEM	3		"Major" comorbidities
532.01	AC DUODEN ULC W HEM-OBST	3		"Major" comorbidities
532.10	AC DUODENAL ULCER W PERF	3		"Major" comorbidities
532.11	AC DUODEN ULC PERF-OBSTR	3		"Major" comorbidities
532.20	AC DUODEN ULC W HEM/PERF	3		"Major" comorbidities
532.21	AC DUOD ULC HEM/PERF-OBS	3		"Major" comorbidities
532.40	CHR DUODEN ULCER W HEM	3		"Major" comorbidities
532.41	CHR DUODEN ULC HEM-OBSTR	3		"Major" comorbidities
532.50	CHR DUODEN ULCER W PERF	3		"Major" comorbidities
532.51	CHR DUODEN ULC PERF-OBST	3		"Major" comorbidities
532.60	CHR DUODEN ULC HEM/PERF	3		"Major" comorbidities
532.61	CHR DUOD ULC HEM/PERF-OB	3		"Major" comorbidities
533.00	AC PEPTIC ULCER W HEMORR	3		"Major" comorbidities
533.01	AC PEPTIC ULC W HEM-OBST	3		"Major" comorbidities
533.10	AC PEPTIC ULCER W PERFOR	3		"Major" comorbidities
533.11	AC PEPTIC ULC W PERF-OBS	3		"Major" comorbidities
533.20	AC PEPTIC ULC W HEM/PERF	3		"Major" comorbidities
533.21	AC PEPT ULC HEM/PERF-OBS	3		"Major" comorbidities
533.40	CHR PEPTIC ULCER W HEM	3		"Major" comorbidities
533.41	CHR PEPTIC ULC W HEM-OBS	3		"Major" comorbidities
533.50	CHR PEPTIC ULCER W PERF	3		"Major" comorbidities
533.51	CHR PEPTIC ULC PERF-OBST	3		"Major" comorbidities
533.60	CHR PEPT ULC W HEM/PERF	3		"Major" comorbidities
533.61	CHR PEPT ULC HEM/PERF-OB	3		"Major" comorbidities
534.00	AC MARGINAL ULCER W HEM	3		"Major" comorbidities
534.01	AC MARGIN ULC W HEM-OBST	3		"Major" comorbidities
534.10	AC MARGINAL ULCER W PERF	3		"Major" comorbidities
534.11	AC MARGIN ULC W PERF-OBS	3		"Major" comorbidities
534.20	AC MARGIN ULC W HEM/PERF	3		"Major" comorbidities

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
534.21	AC MARG ULC HEM/PERF-OBS	3		"Major" comorbidities
534.40	CHR MARGINAL ULCER W HEM	3		"Major" comorbidities
534.41	CHR MARGIN ULC W HEM-OBS	3		"Major" comorbidities
534.50	CHR MARGINAL ULC W PERF	3		"Major" comorbidities
534.51	CHR MARGIN ULC PERF-OBST	3		"Major" comorbidities
534.60	CHR MARGIN ULC HEM/PERF	3		"Major" comorbidities
534.61	CHR MARG ULC HEM/PERF-OB	3		"Major" comorbidities
535.01	ACUTE GASTRITIS W HMRHG	3		"Major" comorbidities
535.11	ATRPH GASTRITIS W HMRHG	3		"Major" comorbidities
535.21	GSTR MCSL HYPRT W HMRG	3		"Major" comorbidities
535.31	ALCHL GSTRITIS W HMRHG	3		"Major" comorbidities
535.41	OTH SPF GASTRT W HMRHG	3		"Major" comorbidities
535.51	GSTR/DDNTS NOS W HMRHG	3		"Major" comorbidities
535.61	DUODENITIS W HMRHG	3		"Major" comorbidities
537.4	GASTRIC/DUODENAL FISTULA	3		"Major" comorbidities
537.83	ANGIO STM/DUDN W HMRHG	3		"Major" comorbidities
540.0	AC APPEND W PERITONITIS	3		"Major" comorbidities
557.0	AC VASC INSUFF INTESTINE	3		"Major" comorbidities
562.02	DVRTCLO SML INT W HMRHG	3		"Major" comorbidities
562.03	DVRTCLI SML INT W HMRHG	3		"Major" comorbidities
562.12	DVRTCLO COLON W HMRHG	3		"Major" comorbidities
562.13	DVRTCLI COLON W HMRHG	3		"Major" comorbidities
567.0	PERITONITIS IN INFEC DIS	3		"Major" comorbidities
567.1	PNEUMOCOCCAL PERITONITIS	3		"Major" comorbidities
567.2	SUPPURAT PERITONITIS NEC	3		"Major" comorbidities
567.8	PERITONITIS NEC	3		"Major" comorbidities
567.9	PERITONITIS NOS	3		"Major" comorbidities
569.60	COLSTOMY/ENTER COMP NOS	3		"Major" comorbidities
569.61	COLOSTY/ENTEROST INFECTN	3		"Major" comorbidities
569.69	COLSTMY/ENTEROS COMP NEC	3		"Major" comorbidities
569.83	PERFORATION OF INTESTINE	3		"Major" comorbidities
569.85	ANGIO INTES W HMRHG	3		"Major" comorbidities
570.	ACUTE NECROSIS OF LIVER	3		"Major" comorbidities
572.0	ABSCESS OF LIVER	3		"Major" comorbidities
572.4	HEPATORENAL SYNDROME	3		"Major" comorbidities
573.4	HEPATIC INFARCTION	3		"Major" comorbidities
575.4	PERFORATION GALLBLADDER	3		"Major" comorbidities
576.3	PERFORATION OF BILE DUCT	3		"Major" comorbidities
577.2	PANCREAT CYST/PSEUDOCYST	3		"Major" comorbidities
579.3	INTEST POSTOP NONABSORB	1		Malnutrition
580.0	AC PROLIFERAT NEPHRITIS	3		"Major" comorbidities
580.4	AC RAPIDLY PROGR NEPHRIT	3		"Major" comorbidities
580.81	AC NEPHRITIS IN OTH DIS	3		"Major" comorbidities
580.89	ACUTE NEPHRITIS NEC	3		"Major" comorbidities
580.9	ACUTE NEPHRITIS NOS	3		"Major" comorbidities

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
583.4	RAPIDLY PROG NEPHRIT NOS	3		"Major" comorbidities
584.5	LOWER NEPHRON NEPHROSIS	3		"Major" comorbidities
584.6	AC RENAL FAIL_ CORT NECR	3		"Major" comorbidities
584.7	AC REN FAIL_ MEDULL NECR	3		"Major" comorbidities
584.8	AC RENAL FAILURE NEC	3		"Major" comorbidities
584.9	ACUTE RENAL FAILURE NOS	3		"Major" comorbidities
590.2	RENAL/PERIRENAL ABSCESS	3		"Major" comorbidities
596.6	BLADDER RUPT_ NONTRAUM	3		"Major" comorbidities
659.30	SEPTICEMIA IN LABOR-UNSP	3		"Major" comorbidities
659.31	SEPTICEM IN LABOR-DELIV	3		"Major" comorbidities
665.00	PRELABOR RUPT UTER-UNSP	3		"Major" comorbidities
665.01	PRELABOR RUPT UTERUS-DEL	3		"Major" comorbidities
665.03	PRELAB RUPT UTER-ANTEPAR	3		"Major" comorbidities
665.10	RUPTURE UTERUS NOS-UNSP	3		"Major" comorbidities
665.11	RUPTURE UTERUS NOS-DELIV	3		"Major" comorbidities
669.10	OBSTETRIC SHOCK-UNSPEC	3	3	"Major" comorbidities
669.11	OBSTETRIC SHOCK-DELIVER	3	3	"Major" comorbidities
669.12	OBSTET SHOCK-DELIV W P/P	3	3	"Major" comorbidities
669.13	OBSTETRIC SHOCK-ANTEPAR	3	3	"Major" comorbidities
669.14	OBSTETRIC SHOCK-POSTPART	3	3	"Major" comorbidities
669.30	AC REN FAIL W DELIV-UNSP	3		"Major" comorbidities
669.32	AC REN FAIL-DELIV W P/P	3		"Major" comorbidities
669.34	AC RENAL FAILURE-POSTPAR	3		"Major" comorbidities
673.00	OB AIR EMBOLISM-UNSPEC	3	1	"Major" comorbidities
673.01	OB AIR EMBOLISM-DELIVER	3	1	"Major" comorbidities
673.02	OB AIR EMBOL-DELIV W P/P	3	1	"Major" comorbidities
673.03	OB AIR EMBOLISM-ANTEPART	3	1	"Major" comorbidities
673.04	OB AIR EMBOLISM-POSTPART	3	1	"Major" comorbidities
673.10	AMNIOTIC EMBOLISM-UNSPEC	3	1	"Major" comorbidities
673.11	AMNIOTIC EMBOLISM-DELIV	3	1	"Major" comorbidities
673.12	AMNIOT EMBOL-DELIV W P/P	3	1	"Major" comorbidities
673.13	AMNIOTIC EMBOL-ANTEPART	3	1	"Major" comorbidities
673.14	AMNIOTIC EMBOL-POSTPART	3	1	"Major" comorbidities
673.20	OB PULM EMBOL NOS-UNSPEC	3	15	"Major" comorbidities
673.22	PULM EMBOL NOS-DEL W P/P	3	15	"Major" comorbidities
673.23	PULM EMBOL NOS-ANTEPART	3	15	"Major" comorbidities
673.24	PULM EMBOL NOS-POSTPART	3	15	"Major" comorbidities
673.30	OB PYEMIC EMBOL-UNSPEC	3	3	"Major" comorbidities
673.31	OB PYEMIC EMBOL-DELIVER	3	3	"Major" comorbidities
673.32	OB PYEM EMBOL-DEL W P/P	3	3	"Major" comorbidities
673.33	OB PYEMIC EMBOL-ANTEPART	3	3	"Major" comorbidities
673.34	OB PYEMIC EMBOL-POSTPART	3	3	"Major" comorbidities
673.80	OB PULMON EMBOL NEC-UNSP	3	15	"Major" comorbidities
673.81	PULMON EMBOL NEC-DELIVER	3	15	"Major" comorbidities
673.82	PULM EMBOL NEC-DEL W P/P	3	15	"Major" comorbidities

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
673.83	PULMON EMBOL NEC-ANTEPAR	3	15	"Major" comorbidities
673.84	PULMON EMBOL NEC-POSTPAR	3	15	"Major" comorbidities
674.00	PUERP CEREBVASC DIS-UNSP	3	1,3	"Major" comorbidities
682.	OTHER CELLULITIS/ABSCESS*	2		Other infections
682.0	CELLULITIS OF FACE	2		Other infections
682.1	CELLULITIS OF NECK	2		Other infections
682.2	CELLULITIS OF TRUNK	2		Other infections
682.3	CELLULITIS OF ARM	2		Other infections
682.4	CELLULITIS OF HAND	2		Other infections
682.5	CELLULITIS OF BUTTOCK	2		Other infections
682.6	CELLULITIS OF LEG	2	10	Other infections
682.7	CELLULITIS OF FOOT	2	10	Other infections
682.8	CELLULITIS_ SITE NEC	2		Other infections
765.01	EXTREME IMMATUR <500G	3		"Major" comorbidities
765.02	EXTREME IMMATUR 500-749G	3		"Major" comorbidities
765.03	EXTREME IMMATUR 750-999G	3		"Major" comorbidities
781.7	TETANY	3	6,	"Major" comorbidities
785.4	GANGRENE	2	10,11	Gangrene
785.51	CARDIOGENIC SHOCK	3	14	"Major" comorbidities
785.59	SHOCK W/O TRAUMA NEC	3		"Major" comorbidities
787.2	DYSPHAGIA	2	1	Dysphagia
799.1	RESPIRATORY ARREST	3	15	"Major" comorbidities
799.4	CACHEXIA	2		Cachexia
933.1	FOREIGN BODY IN LARYNX	1	15	Miscellaneous throat problems
934.1	FOREIGN BODY BRONCHUS	1	15	Miscellaneous throat problems
958.0	AIR EMBOLISM	3	2,3	"Major" comorbidities
958.1	FAT EMBOLISM	3	2,3	"Major" comorbidities
958.5	TRAUMATIC ANURIA	3		"Major" comorbidities
996.02	MALFUNC PROSTH HRT VALVE	3	14	"Major" comorbidities
996.61	REACT-CARDIAC DEV/GRAFT	3	14	"Major" comorbidities
996.62	REACT-OTH VASC DEV/GRAFT	3		"Major" comorbidities
996.63	REACT-NERV SYS DEV/GRAFT	3		"Major" comorbidities
996.66	REACT-INTER JOINT PROST	3	8	"Major" comorbidities
996.67	REACT-OTH INT ORTHO DEV	3	9	"Major" comorbidities
996.69	REACT-INT PROS DEVIC NEC	3	9	"Major" comorbidities
997.62	INFECTION AMPUTAT STUMP	3	9,10,11	"Major" comorbidities
998.0	POSTOPERATIVE SHOCK	3		"Major" comorbidities
998.3	POSTOP WOUND DISRUPTION	3		"Major" comorbidities
998.5	POSTOPERATIVE INFECTION*	2		Other infections
998.51	INFECTED POSTOP SEROMA	2		Other infections
998.59	OTHER POSTOP INFECTION	2		Other infections
998.6	PERSIST POSTOP FISTULA	3		"Major" comorbidities
999.1	AIR EMBOL COMP MED CARE	3	3	"Major" comorbidities

Table 4.10 (cont.)

ICD-9-CM	Name	Tier	RIC Exclusions	Analysis Variable
V440	TRACHEOSTOMY STATUS	1	15	Tracheostomy
V451	RENAL DIALYSIS STATUS	2		Dialysis
V461	DEPENDENCE ON RESPIRATOR	1	15	Ventilator
V4975	STATUS AMPUT BELOW KNEE	3	10	Amputation of LE
V4976	STATUS AMPUT ABOVE KNEE	3	10	Amputation of LE
V4977	STATUS AMPUT HIP	3	10	Amputation of LE

NOTE: Asterisks denote category codes. Additional digits are required for legal codes.

5. UNUSUAL CASES

The classification system discussed in the previous two sections was designed by analyzing more than three-fourths of all rehabilitation cases. However, in its design we deliberately omitted cases that might distort the analysis--i.e., transfer cases, in-hospital deaths, and very-short-stay cases. Here we consider how we should classify these unusual cases.

ISSUES RELATED TO PAYING FOR UNUSUAL CASES

Which Cases Are Candidates for Special Payment

Some cases are transferred to another health care institution before the patient has received the full course of rehabilitation therapy. Many of these transfers are to acute care facilities and occur because the patient has encountered a medical condition or event requiring acute care management. A small number of transfers are from one rehabilitation facility to another, often at the patient's request. There are also a small number of transfers to a long-term care (LTC) hospital. These and transfers to a skilled nursing facility (SNF) indicate that the patient has not recovered adequately to return to independent living. In some cases, the patient is not able to tolerate the intense therapy required in an inpatient rehabilitation setting and the discharge occurs very quickly.

Other patients who are also of concern are those transferred to a sub-acute unit, a nursing home (NH), or a long-term-care home that does not receive Medicare payment. One might be concerned that the purpose of these transfers is to increase reimbursement from non-Medicare sources. However, before acting on this concern, we must define which institutions should be labeled "nursing homes" rather than a form of community living. There is now a continuum of living arrangements--from assisted living through various stages of dependence--and some nursing services are provided at all levels. The decision of many seniors to move into a nursing home or other group residence is greatly influenced by many factors beyond the condition of the patient (and thereby beyond the control of the rehabilitation hospital and its discharge process).

These factors include patient and family preferences, the amount of time that family caregivers are available, and family finances. Thus one cannot use just the presence of nursing services to determine which institutions should qualify as receiving a transfer. CMS has suggested that the right criterion be that the institution qualify to receive either Medicare or Medicaid payment as a nursing home. This is a sensible criterion for defining a transfer if it is possible to distinguish Medicaid payments to a nursing home from Medicaid payments to residential facilities such as group homes for the disabled and assisted living facilities that receive payments under Medicaid Home and Community Waivers.[1]

For a transfer to a hospital, SNF, or Medicaid nursing home, it may be appropriate not to provide the full case payment when the patient receives less than the typical amount of treatment. A reduced payment that reflects the actual cost of the case may be more appropriate for two reasons. It would reduce financial incentives to increase reimbursement by providing care in multiple settings. It also would make funds available that could be used to increase payment for other cases so that their payment would be closer to their costs.

It is worth noting that we do not include transfers to home health agencies (HHAs) or to the outpatient department of the hospital among cases that do not receive the full course of inpatient rehabilitation treatment. Even when patients are functioning well enough to return to the community and no longer need the intense therapy provided in the inpatient rehabilitation setting, they often still need a lesser amount of therapy or medical care, which can be provided efficiently either in the home or in an ambulatory setting. The majority of rehabilitation hospitalizations end with a transfer to home health agencies.[2] Indeed,

---

[1] Several persons have objected that residents of nursing homes are sometimes admitted to (and benefit substantially from) rehabilitation and therefore should be excluded from the transfer rule when they return to such a home from rehabilitation. Although the policy argument is debatable, this problem is extremely infrequent in our data. In both 1998 and 1999, only 1 percent of UDSmr cases were living in a nursing home prior to hospitalization.

[2] Medicare paid for HHA services during the 30 days following discharge for 56 percent of 1996-1997 inpatient rehabilitation patients.

it is much less efficient to keep the patient in the costly inpatient rehabilitation setting than to provide the therapy at an alternative site. The payment systems for the HHA, hospital outpatient department, and Part B therapy should recognize how the needs of patients who are discharged from inpatient rehabilitation differ from the needs of those who are discharged directly from acute care, but these payment systems are outside the bounds of this study.

In studying transfers from rehabilitation hospitals to institutional care, we distinguish those that go to hospitals from those that go to SNFs and other NHs. We use a single hospital group that includes transfers to acute care, to LTC hospitals, and to other rehabilitation hospitals. Transfers to hospitals were almost entirely transfers to acute care (94 percent of hospital transfers in each year). Transfers to other rehabilitation hospitals were very rare--accounting for less than 5 percent of hospital transfers in both 1998-1999 and 1996-1997. LTC hospitals received the small number of remaining transfers to hospitals (1.3 percent of hospital transfers in 1998-1999 and 1.1 percent in 1996-1997). We also use a single group for SNF and other NHs. As discussed earlier in Section 2, our data do not distinguish well between these SNFs and other LTC facilities, so we cannot say with precision which fraction of the transfers to SNFs or nursing homes actually go to SNFs.

In our descriptive statistics for transfer cases, we defined short-stay transfers as those that occur before the (arithmetic) mean length of stay for non-transfer patients in the same FRGC. We chose to use a FRGC-specific cutoff for short stays because of the wide variation across FRGCs in average LOS. We used our recommended definition of FRG from Table 3.16, subdivided by four levels of comorbidity tiers (the three tiers defined in Table 4.10 and a fourth "no relevant comorbidity" tier).

Two other groups of patients are of concern because they have not received the full course of therapy: Those who die in the rehabilitation hospital and those who stay in the hospital a very short time. There are

---

Carter et al.(1997) found that 53 percent of their sample of CY 1994 discharges also used HHA services within 30 days of discharge.

Table 5.1

Transfer Rates from 1996 through 1999,
by Discharge Destination and Type of Rehabilitation Facility

| | | | Percentage of Cases | | | | | |
| | | | All Transfers | | | Short-Stay Transfers | | |
Year	Type of Hospital	No. of Cases	To Hosp.	To SNF/NH	Both Destin-ations	To Hosp.	To SNF/NH	Both Destin-ations
96	All	166,975	7.83	13.10	20.93	6.20	7.80	14.00
97	All	201,212	8.00	13.50	21.50	6.30	8.00	15.20
98	All	227,756	8.19	13.87	22.06	6.33	7.46	13.78
99	All	249,443	8.25	13.17	21.42	6.52	7.24	13.76
96	Free-standing	63,520	8.90	11.60	20.50	6.70	5.50	12.20
97	Free-standing	80,555	9.30	12.40	21.70	6.90	5.60	12.50
98	Free-standing	94,827	8.96	14.32	23.27	6.39	5.60	11.99
99	Free-standing	105,921	9.20	13.30	22.50	6.81	5.23	12.04
96	Unit	103,455	7.10	14.00	21.10	5.80	9.20	15.00
97	Unit	120,657	7.00	14.20	21.20	5.90	9.50	15.40
98	Unit	132,929	7.64	13.56	21.20	6.28	8.78	15.06
99	Unit	143,522	7.55	13.08	20.64	6.32	8.72	15.03

Transfers to hospitals increased modestly in this period--from 7.83 percent of discharges to 8.25 percent. Although transfers to SNFs/NHs increased in the early period, they declined in 1999. The relative decline was larger for short-stay transfers. Short-stay transfers were a smaller percentage of discharges in 1999 than in 1996. This decline is not due to a change in our sample. The rate of transfers to SNFs/NHs declined even within the subset of hospitals that were in our sample in both years.[3]

Characteristics of Transfers

Table 5.1 also shows transfer rates separately for units and freestanding hospitals. Freestanding hospitals were more likely than units to transfer cases to acute care. It may be that units have ready access to consultation from physicians who work in other parts of the

---

[3] Although it would require more analysis to be sure, it is tempting to speculate that the decline was related to difficulty in obtaining a discharge placement in the first years of the SNF PPS.

hospital and can provide preventive services and help assess and monitor rehabilitation patients when they get sick. This may allow the unit to keep an ill patient in a rehabilitation bed. If this interpretation is correct, it reflects the valuing of patient care standards over financial incentives, because the same acute care hospital would receive the acute PPS DRG payment. In the first years of our data, units were more likely to transfer cases to SNFs and nursing homes than were freestanding hospitals. By 1999 the rates of transfer to SNFs and NHs were very similar in both types of facilities. However, units remain more likely to have short-stay transfers to SNFs and NHs.

Table 5.2 provides the transfer rate in our 1998-1999 data by RIC. The patterns by RIC are quite stable and consistent with the 1996 and 1997 data that were presented in Table 4.2 of our interim report. In particular, the large RIC 8, joint replacement of the lower extremity, has very few transfers, which reflects the relative health and strength of this population. On the other hand, RICs 1, 2, 3, 4, and 18 have relatively high transfer rates.

Table 5.2

Transfer Rates by RIC, 1998 and 1999

Year	RIC	No. of Cases	All Transfers			Short-Stay Transfers		
			To Hospital	To SNF/NH	Both Des-tinations	To Hospital	To SNF/NH	Both Des-tinations
98	1	54,777	9.21	20.57	29.79		12.04	19.67
98	2	2,900	11.79	20.52	32.31		11.79	20.97
98	3	5,034	12.02	16.89	28.90	9.56	9.44	18.99
98	4	1,409	13.06	17.32	30.38	9.79	9.44	19.23
98	5	7,077	9.88	11.33	21.21	7.79	6.17	13.96
98	6	10,601	8.62	14.23	22.86	6.36	6.75	13.11
98	7	26,237	6.49	19.50	25.99	4.95	10.81	15.76
98	8	46,886	3.28	5.17	8.45	2.25	2.44	4.69
98	9	10,712	6.84	14.42	21.27	7.63	7.38	12.48
98	10	8,175	12.18	12.45	24.64	9.17	6.48	15.60
98	11	823	16.77	13.73	30.50	12.39	6.68	19.08
98	12	4,884	6.04	9.21	15.25	4.57	3.75	8.31
98	13	2,497	8.01	11.01	19.02	5.89	5.49	11.37
98	14	9,034	11.68	8.52	20.20	9.10	4.01	13.11
98	15	5,899	12.90	9.66	22.56	10.10	4.37	14.48
98	16	3,248	7.14	10.31	17.46	5.02	4.62	9.64
98	17	2,155	7.56	18.33	25.89	5.89	10.30	16.19
98	18	370	16.22	20.00	36.22	12.70	9.73	22.43
98	19	413	11.62	12.59	24.21	10.41	7.26	17.68
98	20	24,465	11.94	12.95	24.88	8.94	6.32	15.25
98	21	157	13.38	13.38	26.75	10.19	5.73	15.92
99	1	54,908	9.02	20.17	29.19	7.61	11.94	19.54
99	2	3,201	12.68	19.43	32.11	10.81	10.78	21.59
99	3	5,523	11.91	16.39	28.30	9.54	9.31	18.85
99	4	1,468	14.51	16.42	30.93	10.69	8.79	19.48
99	5	7,766	9.35	11.13	20.47	7.28	5.82	13.10
99	6	12,150	9.18	14.09	23.27	6.95	6.95	13.91
99	7	28,736	6.84	18.86	25.70	5.40	11.06	16.46
99	8	49,923	3.21	4.76	7.97	2.43	2.22	4.64
99	9	12,455	6.62	14.44	21.05	4.98	7.68	12.66
99	10	8,458	11.74	12.25	23.99	8.87	6.67	15.54
99	11	930	14.52	11.40	25.91	10.86	5.91	16.77
99	12	6,281	6.64	9.78	16.41	5.02	4.51	9.52
99	13	3,061	7.12	12.35	19.47	5.72	5.85	11.56
99	14	10,564	12.07	7.41	19.48	9.61	3.45	13.05
99	15	7,199	13.09	8.20	21.28	10.06	4.15	14.21
99	16	3,829	7.76	9.61	17.37	5.69	4.70	10.39
99	17	2,300	8.13	16.26	24.39	6.61	9.39	16.00
99	18	383	15.40	13.84	29.24	12.79	9.92	22.72
99	19	408	11.27	10.05	21.32	10.54	6.37	16.91
99	20	29,744	11.84	11.72	23.56	9.13	5.90	15.03
99	21	155	15.48	16.13	31.61	12.90	10.32	23.23

Not only do transfers occur disproportionately in certain RICs, they are also highly concentrated in cases with comorbidity and low

function. As shown in Table 5.3, the rates of transfers to hospital are
two and a half times higher for cases in the first comorbidity tier than
for cases with no comorbidity. Transfers to SNFs are also more frequent
in comorbidity tiers 1 and 2 than in tier 3 or for the patients with no
relevant comorbidity. The decline in transfers to SNF/NH from 1998 to
1999 was larger for the patients in the two highest tiers.

Table 5.3

Transfer Rates by Comorbidity Tier

| | | | Percentage of Cases | | | | | |
| | | | All Transfers | | | Short-Stay Transfers | | |
Year	Comorbidity Tier	No. of Cases	To Hospital	To SNF/NH	Both Destinations	To Hospital	To SNF/NH	Both Destinations
98	1	1,616	16.58	18.32	34.90	12.93	9.84	22.77
98	2	14,128	14.52	17.16	31.68	11.47	9.22	20.69
98	3	27,261	13.94	14.67	28.62	11.22	7.67	18.90
98	None	184,751	6.78	13.47	20.25	5.15	7.27	12.42
99	1	1,628	16.65	15.91	32.56	13.70	8.11	21.81
99	2	15,994	14.72	15.44	30.16	11.58	8.51	20.09
99	3	30,235	13.90	14.04	27.94	11.19	7.81	18.99
99	None	201,586	6.82	12.84	19.66	5.37	7.04	12.41

Table 5.4 shows the transfer rates for patients in the lowest- and
highest-numbered FRGs in each RIC. The lowest-numbered FRG generally
contains the patients with the most functional independence; the
highest-numbered FRG contains the patients with the least independence.
Within most RICs, the transfer rates are four to 10 times higher in the
FRGs with the lower-functioning patients than in the FRGs with the
highest-functioning patients.

Table 5.4

Transfer Rates for Lowest- and Highest-Functioning Patients in Each RIC

		Percentage of Cases					
		All Transfers			Short-Stay Transfers		
FRG	No. of Cases	To Hospital	To SNF/NH	Both Desti-nations	To Hospital	To SNF/NH	Both Desti-nations
101	626	2.56	2.24	4.8	2.08	0.64	2.7
114	18,352	16.25	37.54	53.8	14.26	24.72	39.0
201	675	5.93	5.33	11.3	4.44	1.93	6.4
205	1,819	19.13	33.48	52.6	16.44	21.11	37.5
301	2,391	7.15	6.82	14.0	4.77	2.93	7.7
304	1,848	21.48	30.52	52.0	19.16	20.02	39.2
401	456	4.82	6.36	11.2	3.29	3.51	6.8
404	640	24.22	29.38	53.6	19.38	15.47	34.8
501	3,235	4.17	2.81	7.0	2.63	1.17	3.8
505	3,742	17.96	24.21	42.2	15.29	15.02	30.3
601	2,758	3.95	5.00	9.0	2.47	1.70	4.2
604	6,747	13.27	23.98	37.2	10.45	13.15	23.6
701	7,958	3.37	5.92	9.3	2.26	2.53	4.8
705	17,885	10.32	33.10	43.4	8.59	20.58	29.2
801	11,533	1.87	1.40	3.3	1.13	0.59	1.7
806	7,731	9.74	22.71	32.5	7.67	12.35	20.0
901	3,916	3.09	5.59	8.7	1.86	2.32	4.2
904	5,487	11.96	27.76	39.7	9.86	17.17	27.0
1001	1,112	3.69	3.15	6.8	2.43	1.62	4.0
1005	5,047	19.52	22.03	41.5	15.79	13.08	28.9
1101	480	7.50	6.46	14.0	4.17	2.71	6.9
1103	424	24.29	23.35	47.6	19.10	14.15	33.3
1201	1,383	2.31	1.81	4.1	1.52	0.87	2.4
1205	2,287	10.76	23.26	34.0	8.88	10.41	19.3
1301	1,213	4.12	4.95	9.1	2.80	1.81	4.6
1304	1,167	12.68	25.96	38.6	10.45	13.97	24.4
1401	4,646	6.63	2.99	9.6	4.80	1.27	6.1
1404	3,196	21.50	18.90	40.4	18.18	10.08	28.3
1501	2,020	5.25	2.48	7.7	3.91	0.94	4.9
1504	2,152	26.49	19.80	46.3	21.79	11.43	33.2
1601	4,443	5.60	6.10	11.7	3.53	2.39	5.9
1602	2,388	11.73	18.09	29.8	9.38	9.38	18.8
1701	1,374	4.80	7.50	12.3	3.64	2.69	6.3
1703	1,120	12.95	31.79	44.7	11.25	21.34	32.6
1801	70	7.14	8.57	15.7	4.29	2.86	7.1
1804	221	26.24	31.67	57.9	22.62	21.27	43.9
1901	248	6.85	2.82	9.7	5.24	1.21	6.5
1903	244	20.08	21.72	41.8	19.26	14.34	33.6
2001	10,439	6.58	5.31	11.9	4.44	2.05	6.5
2005	6,181	22.28	20.58	42.9	18.72	12.10	30.8
2101	112	9.82	6.25	16.1	6.25	3.57	9.8
2102	196	17.35	19.90	37.2	14.80	10.71	25.5

Resource Use

There are substantial differences in resource use between cases
transferred to a hospital and cases transferred to SNFs and nursing
homes. Table 5.5 shows that the average LOS for cases transferred to a
hospital in 1999 was 12.32 days, whereas the average LOS for cases
transferred to SNFs or NHs was 20.61 days. For comparison, we also show
the average LOS of typical cases (i.e., excluding transfers, in-hospital
deaths, and cases with a LOS of three days or less) with the same
distribution of FRGCs as the transfer cases. The average LOS of non-
transfer cases matched by FRGC to transfers to hospitals was 19.02 days
in 1999. The next column of the table shows the ratio of the LOS of
typical cases to the LOS of similar transfer cases. For the average
transfer to hospital in 1997, a typical case in the same FRGC has a LOS
that is 1.54 times the LOS of the transfer (= 19.02/12.32). On the other
hand, transfers to SNFs have an average LOS similar to that of non-
transfer cases in the same FRGC. As described above, transfer cases tend
to be in FRGCs with higher weights. These FRGCs also have longer LOSs.
The average LOS for non-transfer, non-death cases was only 15.4 in the
1998-1999 period, substantially smaller than the average LOS for non-
transfers with the same FRGC distribution as transfer cases.

Short-stay transfers have shorter LOSs than all transfers, of
course. Typical cases have a LOS that is 1.78 times that of short-stay
transfers in the same FRGC. The LOS of typical cases is more than double
that of comparable short-stay transfers to hospitals and over 50 percent
more than the LOS of short-stay transfers to SNFs and nursing homes.

Table 5.5 also shows the standardized (i.e., wage-adjusted) cost of
each type of transfer and compares it to the standardized cost of
typical cases in the same FRGC. Because transfer cases tend to fall into
more-expensive FRGCs, the average standardized cost of typical patients
with the same FRGC distribution as all transfer cases ($14,892 in CY
1999) is about 32 percent higher than the average standardized cost of
all typical patients ($11,248 in the combined 1998-99 data). If typical
patients were paid an amount equal to the average cost of a case in
their FRGC and transfer cases received the full FRGC payment, the ratio
in the last column would give the payment-to-cost ratio for transfer
cases. Transfers to hospitals would be overpaid by 50 percent, while

transfers to SNFs/NHs would have about the same payment-to-cost (PTC)
ratio as typical cases.

<div align="center">

Table 5.5

**Average LOS and Standardized Cost of Transfer Cases Compared to Those
for Typical Cases in the Same FRGC, by Destination of Transfer Case
and Year**

</div>

Year	Type of Transfer	Number of Cases	Average LOS	LOS of Typical Case in Same FRGC	Ratio	Average Standard Cost	Standard Cost of Typical Case in Same FRGC	Ratio
98	SNF/NH	31,599	20.82	20.12	0.97	15,171	15,237	1.00
98	Hospital	18,653	12.59	19.06	1.51	9,775	14,517	1.49
98	Both	50,252	17.76	19.73	1.11	13,168	14,969	1.14
98	Short-stay SNF/NH	16,985	13.62	20.96	1.54	10,650	15,942	1.50
98	Short-stay hospital	14,408	8.79	19.58	2.23	7,151	14,978	2.09
98	Both short-stay	31,393	11.40	20.33	1.78	9,044	15,499	1.71
99	SNF/NH	32,864	20.61	20.07	0.97	15,530	15,169	0.98
99	Hospital	20,579	12.32	19.02	1.54	9,930	14,449	1.46
99	Both	53,443	17.42	19.67	1.13	13,374	14,892	1.11
99	Short-stay SNF/NH	18,055	13.64	20.78	1.52	11,076	15,769	1.42
99	Short-stay hospital	16,276	8.65	19.39	2.24	7,293	14,769	2.03
99	Both short-stay	34,331	11.27	20.12	1.78	9,282	15,295	1.65

Short-stay transfers are, of course, substantially cheaper than
other cases. If one believed that hospitals would act on the stronger
incentive to discharge early under the PPS than under TEFRA, one might
wish to lower payment for short-stay transfers to SNFs/NHs as well as to
hospitals.

ANALYSES OF THE COST OF TRANSFER CASES

A possible payment system would be to make a single payment rate for all short-stay transfer cases. Because short-stay transfers to SNFs/NHs cost 50 percent more than short-stay transfers to hospitals ($11,076 versus $7,293 in 1999; see Table 5.5), one might have two transfer payment amounts, depending on destination. For either policy, the standard deviation of standardized cost is similar to the standard error of the estimate used in Section 3 and is a reasonable measure of the typical difference between cost and payment under such a policy. For all short-stay transfers, the standard deviation in 1999 was $5,833, or 63 percent of the mean. For short-stay transfers to hospitals only, the standard deviation in 1999 was $5,263, or 72 percent of the mean; for SNF/NHs, it was $5,739, or 52 percent of the mean. So any system with only one or two standardized payments for transfer cases would allow substantial overpayment for many cases and underpayment for others. Further, for patients who are near the average LOS, a large increase in payment could be obtained for a small increase in resources, leading to an incentive to game the system. A smoother change in payment with LOS would lower the incentive to manipulate care for financial reasons.

How the Costs of Short-Stay Transfers Compare to Those of Typical Cases

Although part of the reason for variation in standardized costs across transfer cases is the FRGC of the case, the larger reason is the variation in LOS. The inter-quartile range of LOS for short-stay transfers is from seven to 15 days. Almost 10 percent of cases have a LOS of three days or less, and 10 percent have a LOS of 20 days or more. Consequently, we have explored the relationship between the cost of short-stay transfers and per diem cost in the FRGC.

We used simple regression models that capture the major variation across both LOS and FRGC group, and, if necessary, transfer destination. This analysis is similar to the payment model developed in our earlier work on the model rehabilitation PPS that, in turn, built on earlier analyses of the costs of transfer cases in the acute care PPS (Carter and Rumpel, 1993). The models build on the fact that the costs of hospitalizations are well modeled as per diem cost times LOS plus an

extra cost for the first day to cover costs that occur only once in the stay.

We present two models. In Model 1 we show how the per diem cost of transfer cases relates to the per diem cost of typical cases. In particular, we regress the standardized cost of each transfer case on two variables: (1) the average standardized cost of a typical day in the FRGC or the per diem cost of the FRGC (average standardized cost of a non-transfer, non-death case in the FRGC divided by the average LOS for the same set of cases); (2) per diem cost for all days beyond the first--i.e., the LOS of the transfer case minus 1, multiplied by the per diem cost. The model has no intercept term, so per-case costs, if any, are loaded onto the first day per diem.

Model 2 tests whether transfers to hospitals have substantially different costs from transfers to SNFs and nursing homes. It does so by adding the interaction of the two variables in Model 1 with a dummy variable that is 1 if and only if the transfer was to a hospital (rather than to a nursing facility).

In each year and in each model, the cost of the first day of each stay appears to be a little more than one and a half times the cost of a typical day in the same FRGC (see Table 5.6). The coefficient on the per diem pay is very close to 1. The standard error of the estimate of the coefficient on the cost of subsequent days is approximately 0.0034 in each year, and therefore the coefficient is statistically indistinguishable from 1 ($t < 1$, $p > 0.2$) in 1996 and 1997, slightly less than 1 in 1998, and slightly more than 1 in 1999. So additional days in the transfer stay cause an average increase in cost very close to the average cost of a typical day in the FRGC.

Table 5.6

**Regressions of Cost of Short-Stay Transfer Cases
on Per Diem Payment and First Day Payment**

	1996		1997		1998		1999	
	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.
Model 1								
First day cost	1.630	35.82	1.65	38.58	1.643	37.93	1.743	40.79
Subsequent day cost	0.997	285.01	0.998	292.87	0.987	279.86	1.023	290.49
Model 2								
First day cost	1.688	22.31	1.67	23.60	1.610	21.91	1.769	23.61
Subsequent day cost	0.991	194.46	0.989	200.67	0.984	189.04	1.017	191.17
Transfer to hospital and:								
First day cost	-0.129	1.34	-0.138	1.53	-0.147	0.16	-0.089	0.97
Subsequent day cost	0.018	2.35	0.038	5.04	0.020	2.60	0.021	2.66

Note: R-squared for Model 1, by year: 0.791, 0.764, 0.894, 0.893.
R-squared for Model 2, by year: 0.791, 0.765, 0.894 and 0.893.

The results from Model 2 show that, after controlling for FRGC and LOS, the cost of cases transferred to hospitals is very similar to the cost of cases transferred to SNFs. The coefficient on first day cost interacted with transfer to a hospital is not significant in any of the four years. Although the coefficient on the interaction of hospital transfer and subsequent per diem is significant, the magnitude of the effect is very small and the effect is in an opposite direction to the effect of the constant term. Thus we believe that a single per diem payment rule could closely approximate cost for both types of transfers.

The note in Table 5.6 gives the R-squared from each of the models. The standard errors of the estimate (SEEs) are roughly one-third of the standard deviation of the cost of short-stay transfer cases (see/std = sqrt(1 - R-squared)). Thus, using a per diem payment rule greatly improves accuracy compared to either using a single payment for transfer cases or using payments that depend only on transfer destination.

In the interim report, we used this equation to recommend that HCFA pay a per diem equal to the average payment for a day in the FRGC plus

an additional one-half of a per diem. We also noted that eliminating the one-half of a per diem would provide a small penalty for each transfer case, which might discourage hospitals from developing a care policy based on frequent transfers but have only a small effect on access. In the NPRM, HCFA proposed paying short-stay transfers a per diem equal to the average payment for a day in the FRGC.

How the Definition of "Typical" Affects the Transfer Cost Model

The recommendation in our interim report did not take into account that the average payment for a day in the FRGC may vary from the average daily cost for a typical case. The typical cost used in the weight calculation, which produces the numerator of the per diem payment, includes the cost of long-stay transfers--but these were not included in the cost used in the regressions in Table 5.6. Further, the per diem payment weight excludes cases with a LOS of three days or less, which were used in this regression. Further, the average LOS used to divide the case payment amount was based on the average LOS of cases discharged to the community and thus did not match the average LOS of the cases in the weight calculation.

The result of these inconsistencies is shown in the second regression in Table 5.7, which shows how the cost of short-stay transfers varies with the cost of the cases that are paid as typical divided by the LOS of truly typical cases. Thus this regression shows how costs vary with the per diems in our recommendation. To make comparison easier, the first regression in Table 5.7 is Model 1 from Table 5.6. Transfer case costs are not directly proportional to the LOS times the per diem plus a case cost. Using the NPRM rule of just LOS times a per diem, very short transfer cases are underpaid, but longer short-stay transfer cases are overpaid relative to costs.

Table 5.7

**Regressions of Cost of Short-Stay Transfer Cases on Daily Cost of
Typical Cases, for Various Definitions of Typical Case**

Definition of Typical Case	Variable	1998		1999	
		Coef.	t-stat.	Coef.	t-stat.
Discharged to community					
	First day cost	1.643	37.93	1.743	40.79
	Each subsequent day cost	0.987	279.86	1.023	240.49
Add long-stay transfers and remove short-stay					
	First day cost	1.657	41.07	1.755	44.08
	Each subsequent day cost	0.897	274.90	0.929	285.13
Fix typical average LOS					
	First day cost	1.610	36.80	1.709	39.61
	Each subsequent day cost	1.003	280.95	1.039	291.73

Note: R-squared for 1998 are, in model order: 0.894, 0.892, 0.894.
R-squared for 1999 are, in model order: 0.894, 0.892, 0.894.

The problem with this model lies in the inconsistency between the weight calculation (which includes long-stay transfers) and the average LOS (which does not include long-stay transfers). This can be fixed if we use all cases paid as typical cases in the FRGC as both numerator and denominator in the per diem calculation. This is shown in the last regression in Table 5.7. We find that the coefficient on subsequent days has returned to 1.

It might appear circular to use long-stay transfers to define the LOS cutoff for long-stay transfers, but in fact small changes in the cutoff value cause even smaller changes in the average LOS--since the cases being included or excluded are close to the average LOS by definition. Consequently, the process will converge after a very few iterations.

VERY-SHORT-STAY CASES

Another group of patients who are unlikely to receive the full course of rehabilitation are those who stay in the hospital a very short time. In Table 5.8, we show the number of cases with a LOS up to five

Table 5.8

**Number and Percentage of Cases With and Without FIM Assessment
and With Each Discharge Destination, by LOS (1998 and 1999 data)**

	Number of Cases						
	LOS=1	LOS=2	LOS=3	LOS=4	LOS=5	LOS=6	All
FIM known							
SNF/NH	58	175	390	597	642	62,562	64,424
Hospital	1,627	1,934	2,120	2,279	2,247	28,994	39,201
AMA	51	74	82	80	68	730	1,085
Died	70	127	151	167	141	1,928	2,584
Community	589	2,182	5,622	10,544	12,580	338,191	369,708
Subtotal	2,395	4,492	8,365	13,667	15,678	432,405	477,002
FIM missing							
SNF/NH	7	9	2	0	1	6	25
Hospital	153	87	45	13	4	16	318
AMA	17	6	1	0	0	0	24
Died	9	4	2	0	0	2	17
Community	78	39	40	13	4	130	304
Subtotal	264	145	90	26	9	154	688
Total	2,659	4,637	8,455	13,693	15,687	432,559	477,690
	Percentage of Cases						
	LOS=1	LOS=2	LOS=3	LOS=4	LOS=5	LOS=6	All
FIM known							
SNF/NH	2.18	3.77	4.61	4.36	4.09	14.46	13.49
Hospital	61.19	41.71	25.07	16.64	14.32	6.70	8.21
AMA	1.92	1.60	0.97	0.58	0.43	0.17	0.23
Died	2.63	2.74	1.79	1.22	0.90	0.45	0.54
Community	22.15	47.06	66.49	77.00	80.19	78.18	77.39
Subtotal	90.07	96.87	98.94	99.81	99.94	99.96	99.86
FIM missing							
SNF/NH	0.26	0.19	0.02	0.00	0.01	0.00	0.01
Hospital	5.75	1.88	0.53	0.09	0.03	0.00	0.07
AMA	0.64	0.13	0.01	0.00	0.00	0.00	0.01
Died	0.34	0.09	0.02	0.00	0.00	0.00	0.00
Community	2.93	0.84	0.47	0.09	0.03	0.03	0.06
Subtotal	9.93	3.13	1.06	0.19	0.06	0.04	0.14
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Note: AMA = against medical advice.

days and then group all cases with a LOS of six days or more. To
describe these cases further, we grouped them by information about their
discharge destination. Because of the small number of cases per cell, we
added the data for both years together. The first half of the table
covers our analysis sample, for which we had good cost and FIM data and
could assign an FRGC. The second half of the table includes the small

number of patients who left the rehabilitation hospital prior to completion of their initial assessment. UDSmr (but not Caredata.com or HealthSouth) indicates these cases in their admission class data field with a 4 for "unplanned discharge without assessment." According to the coding instructions, "This is a stay that lasts less than 72 hours because of an unplanned discharge. An admission FIM assessment has not been completed." All the UDSmr cases with missing FIM data were classified as "unplanned discharge without assessment," although many stayed more than three days. Although Caredata.com and HealthSouth did not explicitly classify patients in this manner, each source did have cases that were not assigned answers to all the FIM questions.[4]

Two-thirds of the cases with only a one-day stay were transferred to the hospital, and 2 to 3 percent each went to a nursing home, died, or left against medical advice. However, about 25 percent (= 22.15 + 2.93) were discharged to the community. The fraction of patients in each LOS who are discharged to the community rises quickly with LOS and is roughly similar for all cases with LOS greater than 3. Thus, although we can distinguish the reason why some patients have very short stays, there are other short stays whose cause is not available from our data. But it seems unlikely that LOS of two or three days can be viewed as a typical course of inpatient rehabilitation, even when the patient returns home.

We next asked how much very-short-stay cases cost and how this varied by LOS and by the apparent reason for the very-short-stay case. We restricted the analysis to patients who stayed three days or less. Because we had already analyzed transfers, we also excluded those patients from the data in Table 5.9. We grouped "left against medical advice" with those discharged to the community, because we believe that this distinction may be ambiguous in some cases. These short-stay cases have an average standardized cost of only $2,146, or 19.1 percent of the cost of a typical case that is discharged home.

---

[4] Insofar as cases with missing FIM data did not actually miss assessment, we may have overestimated the extent of the missing assessment problem. However, we believe the overestimate is modest, because 80 percent of the data came from UDSmr and these hospitals explicitly noted the lack of assessment.

Table 5.9

Average and Standard Deviation of Standardized Cost of Selected Very-Short-Stay Cases (combined 1998-1999 data)

	LOS=1			LOS=2			LOS=3			All LOS <=3		
	No. of Cases	Avg. Cost ($)	Std. Dev. Cost	No. of Cases	Avg. Cost ($)	Std. Dev. Cost	No. of Cases	Avg. Cost ($)	Std. Dev. Cost	No. of Cases	Avg. Cost ($)	Std. Dev. Cost
Discharged prior to assessment	264	970	486	145	1,877	901	90	2,620	976	499	1,531	788
Discharged to home or AMA	638	899	602	2,253	1,716	705	5,702	2,514	1,169	8,593	2,185	825
Died	80	1,112	770	131	1,857	642	153	2,717	993	364	2,055	801
Any of above	982	935	619	2,529	1,733	749	5,945	2,521	1,046	9,456	2,146	805

Within each LOS, the three groups have similar costs. Because of the different distribution of LOS, those discharged prior to assessment cost somewhat less than those discharged to home. Although the death cases also cost less than those discharged to home, the difference in cost is only $130 per case. Based on similar data for 1996-1997 cases, we recommended that HCFA create a case mix group for cases with LOS less than or equal to three days, and HCFA proposed using such a CMG in the NPRM.

IN-HOSPITAL DEATH CASES

Table 5.10 shows the death rate by RIC. Although the death rate is low in all RICs, it is very low in RIC 8, joint replacement of the LE. The death rate is also noticeably lower than average in the other orthopedic RICs (7 and 9) and in the pain RIC (16). Although we had no deaths in the burn RIC (21), the sample size is too small to ensure that the underlying death rate is lower than average. Patients in RIC 15 (pulmonary disorders) have the highest death rate.

The patients who died in the hospital had significantly shorter LOSs and lower costs than those in the same FRGC that would be paid as typical cases (i.e., cases discharged to the community and long-stay transfer cases). See the first row of Table 5.11. If each death case

were to receive a payment equal to that of typical cases, the death cases would be overpaid by 51 percent. In the previous subsection, we found that the cost of death cases with LOS of three days or less was very similar to the cost of other cases with similar short stays. Thus, such stays were included within the atypical low-cost CMG in the NPRM. As shown in the second row of Table 5.11, after removing the death cases with LOS <=3, the remaining death cases still cost about one-third more than typical cases.

The size of the overpayment suggests it is desirable to provide a special payment for death cases. One possibility would be to pay each death case the same amount, with the amount set to be proportional to the cost of the average death case. This would result in fair payment for these cases and free up some money to be paid for other cases. As before, we use the standard deviation of standardized cost as a proxy measure for the typical difference between cost and payment under such a policy. As shown in Table 5.12, the standard deviation is $10,301, or 97 percent of the mean cost. Clearly such a single payment would result in payments substantially above or below cost for many cases. As is shown in the next two lines of the table, death cases in the orthopedic RIC have a shorter LOS and 20 percent lower costs than death cases in the non-orthopedic RICs. Setting two payment rates rather than one would decrease the average overpayment of orthopedic cases that die in the hospital, but would have only a little effect on the degree of overpayment and underpayment as measured by the root mean square error, whether or not very-short-stay death cases are included.

Table 5.10

Death Rate by RIC (combined 1998-1999 data)

RIC	Number of Cases	Death Rate (%)
1	109,789	0.63
2	6,111	0.75
3	10,567	0.87
4	2,880	0.73
5	14,853	0.38
6	22,762	0.55
7	55,006	0.36
8	96,838	0.07
9	23,174	0.32
10	16,644	0.91
11	1,753	0.68
12	11,169	0.29
13	5,564	0.47
14	19,621	0.83
15	13,114	1.64
16	7,078	0.32
17	4,460	0.43
18	753	0.66
19	822	0.36
20	54,271	1.06
21	312	0.00
Total	477,541	0.54

Note: Omits only cases where RIC could not be assigned.

Table 5.11

LOS and Standardized Cost of In-Hospital Deaths Compared to Those of Cases in the Same FRGC That Are Paid as Typical Cases (combined 1998-1999 data)

Type of Death Case	Number of Cases	Average LOS	LOS of Typical Case in Same FRGC	Ratio	Average Cost	Cost of Typical Case in Same FRGC	Ratio
All	2,584	13.37	21.71	1.62	10,531	15,865	1.51
Exclude LOS <=3	2,236	15.11	21.67	1.43	11,847	15,840	1.34

Note: Typical cases include cases discharged to the community and long-stay transfers.

Table 5.12

Standardized Cost of In-Hospital Deaths for Selected Partitions of Death Cases (combined 1998-1999 data)

Type of Death Case	Number of Cases	Avg. Cost	Std. Dev. Cost	Group RMSE	Ratio of RMSE to Avg. Cost
All	2,593	10,570	10,301	10,301	0.97
Orthopedic RIC	344	8,461	6,940	10,269	0.97
Non-Orthopedic RIC	2,249	10,892	10,687		
All LOS>3	2,245	11,886	10,464	10,464	0.88
Orthopedic RIC, LOS>3	291	9,645	6,906	10,269	0.86
Non-Orthopedic RIC, LOS>3	1,954	12,219	10,857		
Orthopedic RIC, 3<LOS<=13	192	6,136	2,878	8,229	0.69
Non-Orthopedic RIC, 3<LOS<=15	1,258	7,305	3,709		
Orthopedic RIC, LOS>=14	99	16,450	7,354		
Non-Orthopedic RIC, LOS>=16	696	21,102	13,550		

The last partition in Table 5.12 divides both orthopedic death cases and non-orthopedic death cases into shorter-stay and longer-stay cases based on the average LOS of the relevant death. Payment under such a system, with the omission of death cases with LOS of three days or less, would more accurately match payments to cost with a group RMSE of only 69 percent of the mean. The CMGs in the NPRM included four groups of death cases similar to these four--the LOS cutoffs were updated here to reflect the 1999 data.

INTERRUPTED STAYS

In the NPRM, HCFA proposed that all discharges where the patient returned to the hospital on the day of discharge or either of the following two days be bundled with the following discharge for the purposes of case classification and payment. Alternatives that might be considered include using a different length of interruption as the bundling criterion and bundling the cost of non-physician care during the interruption with the IRF PPS payment.

In this subsection, we address several analysis questions that could shed light on the relative value of the proposed policy compared to alternative policies:

- What are frequent causes of interruptions?

- How much of the cost of acute care during the interruption is built into the TEFRA payment?
- How much do interrupted stays cost relative to payment amounts under alternative policies?

To address these questions, we primarily used the MEDPAR data set and explicitly assumed an interruption to be synonymous with a discharge. However, as described below, we evaluated that assumption based on the UDSmr definition of an interrupted stay.

Data and Methods

To analyze bundling policies, we sorted all MEDPAR discharges from rehabilitation facilities (including those for which we did not have a matched FIM-MEDPAR sample) on provider number, beneficiary identifier, and discharge and admission dates. We then bundled discharges whenever the patient returned to the same rehabilitation facility within two calendar days of discharge.[5] There were a total of 4,325 bundles involving at least one matched FIM-MEDPAR record. These bundles included 8,809 discharges.

We used all these bundles in the analyses of the causes of interruptions and the acute care payments by Medicare during the interruptions. There were 4,484 (= 8,809 - 4,325) interruptions in these bundles. We attempted to find a non-rehabilitation MEDPAR record coinciding with each of the interruptions. We began by finding the non-rehabilitation MEDPAR record for the same beneficiary that immediately preceded each rehabilitation discharge following an interruption. In a few cases, the non-rehabilitation record matched two interruptions in the same bundle and we assigned it to the interruption it matched more closely in time. Finally, we restricted the acute sample to discharges that occurred within one day of the admission date of the rehabilitation discharge following the interruption.[6] We found the acute care record

---

[5] There were 13 discharges in which the patients appeared to have been readmitted to rehabilitation before they were discharged from the same rehabilitation hospital. They were eliminated from the analysis as bad data.

[6] All but eight occurred on the date of rehab admission. We also had data on 1998, but not 1999, stays in SNFs. There were four 1998 SNF stays that apparently matched an interruption. Thus, although we do not

for 3,069 interruptions, or 68.4 percent of all interruptions. The rate at which we found acute stays, however, varied with the length of the interruption. We found stays for 90.5 percent of interruptions that lasted two days, but for only 5 percent of interruptions where the patient returned on the day of discharge and for 63.9 percent of interruptions where the patient returned on the next day. Because of minor problems with dates of admission and discharge, there are likely to be slightly more interruptions with coinciding acute stays than we found.

To simulate bundled payment, we classified the bundles as we had in the interim report. We used the impairment code and FIM scores from the FIM record that matched the first discharge. When there was no FIM record matching the first discharge and the FIM record that matched the second discharge recorded an admission within one day of the first discharge's admission date, we used those FIM data. This occurred in 304 bundles, or 7 percent of all interrupted stays with any match. In many cases, it appeared that the FIM record covered multiple discharges (i.e., the admission date on the FIM record matched the admission date on the first MEDPAR discharge and the discharge date on the FIM record matched the discharge date on the last discharge--yet the FIM record did not note an interrupted stay). We used the last discharge in the interrupted stay to code deaths and the bundle's transfer status. We set the comorbidity tier for the bundle as the highest comorbidity tier that appeared on any matched MEDPAR discharge in the stay. Finally, we summed days and cost across each discharge to assign the LOS and cost for the interrupted stays.

Of the bundles, only 3,118 began with a matched record and therefore could be assigned an FRGC. After eliminating hospitals without good cost data and the few records, if any, without good FIM data, we were down to 3,021 bundles in our sample for analysis of payment issues.

To get some insight into the frequency with which interruptions to the rehabilitation program occur that do not result in a discharge, we

include the costs that Medicare paid to SNFs in our estimate of Medicare payments for nonphysician care within the interruption, we expect that the effect of this omission is less than one-tenth of one percent (because SNF stays are paid less than hospital stays).

used the UDSmr information on interrupted stays. UDSmr defines a stay as interrupted when the patient is transferred to acute care and returns to the same rehabilitation program within 30 days for treatment of the same impairment. For interrupted stays, UDSmr records the starting and ending dates for up to three interruptions. There were a total of 10,639 UDSmr records for interrupted stays that were matched to at least one MEDPAR record. The total number of MEDPAR records matched to these stays was 20,011, or slightly less than two records per stay.

Overall, for 81.8 percent of the interruptions identified on the UDSmr file, we found both TEFRA discharges. Further, the only substantial effect of the length of the interruption is for cases where the patient returned on the day that he or she left the rehabilitation program. For such interruptions, we found both discharges in only 28.6 percent of cases. Thus we must assume that there are more cases with interruptions of less than a day than are captured by our analyses based on discharges alone. However, when the interruption was only one day, we found both discharges in 81.2 percent of cases, slightly more than for the interruptions that lasted from six to 30 days. For interruptions of two days, we found both discharges in 84.7 percent of cases. Thus, excluding the discharges of less than a day, we found both TEFRA discharges for 83.2 percent of bundled discharges. Further, we know there are some errors of omission in our matching algorithm. Thus we expect that some of the missing matching discharges actually occurred and that their omission from our sample is essentially a random event. Since we are missing so few interruptions as classified by UDSmr, we believe that our analyses of the TEFRA discharges will adequately describe the 1998 and 1999 cases that would have been candidates for the bundling policy. Also, we expect that the UDSmr definition of interrupted stays, which is used widely in the field, is consistent with a definition based on multiple discharges except for cases that return to rehabilitation on the day of discharge.

The Acute Episode

Table 5.13 lists the 10 most frequent DRGs that occurred during these very short interruptions. The first seven are all in the cardiovascular Major Diagnostic Category (MDC), and only one of these is surgical (DRG 116). Given that all these patients were discharged within two days, it is not surprising that they are all among the less expensive cardiac DRGs. Cardiovascular MDC DRGs accounted for 37 percent of the interruptions, although cardiac rehabilitation accounted for only 3.0 percent of bundles. The other DRGs in the top 10 are stroke, G.I. hemorrhage, and extracranial vascular procedures. Few of these DRGs, if any, look like they are conditions for which care could have been provided during the initial hospitalization but was postponed.

Medicare paid an average of $7,688 for acute care in the 2,202 bundles for which we found an acute stay. Averaging in the 819 bundles for which we did not find an acute stay (mostly one-day interruptions) by assuming that Medicare paid nothing for acute care during the interruption, the cost of acute care during these interruptions averaged $5,604. We expect there were some stays (and possibly outpatient care), so we believe the acute care cost more than $5,604 per bundle. Because the interruptions for which we did not find acute care hospitalizations were shorter, we believe that the $7,688 is an upper bound on the cost of acute care during interrupted stays where the patient returns to his rehabilitation program within three days.

Table 5.13

DRGs Assigned to Care Between Bundled Discharges

	DRG Number	Frequency
Cardiac arrhythmia and conduction disorders, w CC	138	204
Chest pain	143	174
Cardiac arrhythmia and conduction disorders, without CC	139	107
Atherosclerosis with CC	132	94
Other permanent cardiac pacemaker implant or AICD lead or generator proc.	116	87
Other circulatory system diagnoses with CC	144	86
Heart failure and shock	127	85
Specific cerebrovascular disorders except TIA	14	81
GI hemorrhage w CC	174	76
Extracranial vascular procedures	5	74

Notes: AICD = Automated cardioverter defibrillator.
TIA = transient ischemic attack. GI = gastrointestinal.

The Cost of Bundled Cases

The first line of Table 5.14 shows the resources used by the 3,021 bundled cases with all necessary data to simulate payment. The inpatient rehabilitation cost is substantially more than that for the average discharge. Patients with bundled discharges spent an average of 24.9 days in rehabilitation and cost over \$20,000 per bundled case--almost twice the average 1999 discharge cost of \$11,386. The column labeled discharge payment shows the payment that the case would receive if each of its discharges were paid under the following assumptions: (1) the amount of payment for cases discharged to the community and long-stay transfers is set at the average cost of such cases in the same CMG; (2) payment for a discharge where the patient died and for very short stays is the average cost of a case in the appropriate CMG; and (3) payment for short-stay transfers is a per diem payment equal to the average daily cost in the CMG plus a case payment of one-half of a per diem. The column labeled bundled payment counts only a single payment for the

Table 5.14

LOS, Standardized Cost, and Payment, Under Bundled and Non-bundled
Policies, as a Function of Length of Interruption
(combined 1998-1999 data)

						PTC Ratio	
Length of Interruption	Number of Bundles	Total LOS	Total Cost	Dis-charge Payment	Bundled Payment	Dis-charge Payment	Bundled Payment
0 to 2	3,021	24.9	20,059	20,078	14,116	1.0010	0.7037
0 to 10	10,572	26.1	20,276	19,897	13,683	0.9813	0.6748
0	49	20.9	17,274	19,110	12,878	1.1063	0.7455
1	871	25.6	20,851	19,951	13,460	0.9568	0.6455
2	1,217	25.0	19,890	19,748	13,489	0.9929	0.6782
3	1,492	25.8	20,002	19,743	13,556	0.9871	0.6777
4	1,526	25.5	19,751	19,486	13,511	0.9866	0.6841
5	1,366	25.7	19,918	19,364	13,577	0.9722	0.6817
6	1,233	25.9	19,933	19,586	13,649	0.9826	0.6848
7	1,012	27.0	20,831	20,404	13,913	0.9795	0.6679
8	766	27.7	21,114	20,629	14,138	0.9770	0.6696
9	592	28.4	21,421	20,999	14,235	0.9803	0.6645
10	448	27.6	21,072	20,842	14,139	0.9891	0.6710

Note: Except for the first row of the table, bundles are restricted to
the first two discharges.

entire case, using exactly the same rule and the CMG assigned to the
bundle as described in the Methods section, above. However, the rate of
all bundled payments was increased to reflect savings from the lower
level of payments to bundles. Naturally, most of these savings go to the
non-bundled cases, and thus payments increase by only 0.3 percent.

The per-discharge payment pays these cases almost exactly in
proportion to their cost. The bundled payment pays these cases on
average only 70 percent of costs. If our lower estimate for the cost of
acute care is entered as part of the cost to be borne by the
rehabilitation hospital, the average PTC would decline to 0.55.

We next briefly examine the effect of changing the duration of the
interruption that would trigger bundling. We identified all discharges
in our sample hospitals where the patient returned to rehabilitation at
the same hospital within 10 days. In 2.6 percent of all discharges,
patients returned to their rehabilitation program within 10 days (this
is exactly the same percentage as for the 1996-1997 data covered in our
interim report). Eighty-nine percent of these returns had been
discharged as transfers to a hospital, 9.5 percent were discharged home,

and 1.5 percent were discharged to a SNF or NH. Twenty-eight percent of all transfers to hospitals returned to the same rehabilitation hospital within 10 days to continue their rehabilitation.

We analyzed the cost of interrupted stays as a function of the length of the rehabilitation program interruption. We also compared that cost to payment under both a bundled payment system and one that would pay separately for each discharge. For simplicity, we considered only the first two rehabilitation discharges for each interrupted stay. We used only stays for which both discharges were in our analysis sample. The results are also found in Table 5.14. The second row covers all the interrupted stays where patients returned within 10 days and for which we have data on both parts of the stay. The average patient spent 26.1 days in rehabilitation in the first two parts of the interrupted stay, and that rehabilitation cost an average of $20,276.

We then used the same payment rules that we used to look at payments for the NPRM policy. Paying all these bundles using a per discharge system results in payment that is very close to cost (98.13 percent), whereas a bundled policy pays only about two-thirds of cost (67.48 percent). Looking at the rows of the table, we see that the per discharge policy overpays the cases that return on the day of discharge by 10 percent, while the bundled policy underpays them by 25 percent. After that, there is no strong pattern by LOS--the per discharge pattern pays between 96 and 99 percent of costs, and the bundled policy pays between 65 and 68 percent of costs. Thus, only the cases returning on the day of discharge are overpaid by a per-discharge policy, and the PTC ratio provides no support for choosing any bundling interval beyond the same day level.

Another question is how risky it would be for hospitals to bundle into the payment the acute care costs that are now paid for separately. Medicare paid an average of $7,688 for acute care hospitalizations in the 2,202 bundles for which we found an acute stay. If these costs are bundled and a single IRF payment is made, it would cover only an average

of 51.15 percent of the costs of bundles with acute care.[7] The average
bundled case, including those which did not have (or for which we did
not find) an acute stay would have a PTC ratio of about 55 percent.
(This is an upper bound, because the matching process may have missed
some acute stays--but the expected value is no lower than 51.15
percent.)

IMPLICATIONS FOR PAYMENT

Payment for Transfer Cases

If short-stay transfer cases were to receive a full FRGC payment,
most of them would be substantially overpaid relative to typical cases.
The average PTC ratio for short-stay transfer cases would be about 1.7.
The PTC ratio for transfers to hospitals would be 2.0.

The potential profit from these cases might provide an incentive
for abuse, with some hospitals transferring cases to reduce their costs
and increase payments. Also, some cases might be admitted to
rehabilitation even when there was a high likelihood that the patient
would not be able to sustain the therapy and would need to return to
acute care. To avoid these incentives, it seems prudent to implement a
reduced payment for short-stay transfers, defined as cases
transferred--before the mean LOS for their FRGC--to hospitals, SNFs, or
nursing homes that receive Medicaid payment (assuming the latter
institutions can be clearly separated from all forms of community
living).

It would not be prudent to reduce the payment for such cases much
below expected cost. We would not want to discourage appropriate
transfers that provide beneficiaries with needed services. Further, too-
low payments for transfer cases or very-short-stay cases might have the
unwanted effect of reducing access to inpatient rehabilitation for
patients with a good, but not certain, chance to complete inpatient
rehabilitation and return to the community.

---

[7] This reflects an increased payment rate for the acute care that
HCFA would no longer pay for separately. However, most of the increase
goes to non-bundled cases.

Choosing an appropriate payment policy for short-stay cases and transfers thus requires balancing a reduction in incentive to provide unneeded care or to obtain unjustified payments against the need to provide adequate funding for, and access to, all appropriate care. Using a per diem payment system for transfer cases would closely match payments to cost. Indeed, it would reduce the error in payment by more than half compared to a single per-case transfer payment. This greater accuracy would reduce opportunities for profitable gaming and be fairer to facilities that, just by chance, receive many patients that require a transfer.

Based on our model of the cost of transfer cases, the payment system that would best match payments to cost would be a per diem payment equal to the average payment for a day in the FRGC plus an additional case-level payment of one-half of the per diem. We recommended such a payment to CMS.

Because marginal cost can be less than average cost, it is important that the rate of transfer cases be monitored on a hospital-specific basis to ensure that gaming does not occur.

Payment for Deaths and Atypical Short-Stay Cases

In our matched data set, 1.9 percent of discharges are not transfers and yet stay less than three days. Few, if any, of these cases receive a full course of rehabilitation. Consequently, it makes sense to reduce their payment so that it is based on relative cost. This can be done by creating a special group for these cases. The variance in the cost of all cases with LOS <= 3 days is only 37.5 percent of the mean, so there is no need to split these cases further. However, the creation of a group for such cases provides a strong incentive to keep the patient for four days. This split may result in a decrease in the number of cases with LOS <=3 and a corresponding increase in the number of cases with LOS of four or five days. This should be monitored at both the hospital and national levels.

Even fewer cases die in the hospital--about one-half of one percent. On average these cases cost substantially less than typical cases in the same FRGC, so it is reasonable to pay these cases closer to costs, thereby freeing up funds to pay other cases more accurately.

There is a large variation in the cost of such cases. Splitting the
death group by whether or not the patient was in an orthopedic RIC and
whether or not the patient stayed longer than the average death case
would further increase the accuracy of the payment system. Because of
the large variation in relative costs across FRGCs, any grouping of
death cases into a small number of death-only categories will result in
some cases that are paid more if the patient dies in the hospital than
if the same patient is discharged alive.

Bundling

The analysis above addresses the results of bundling two
rehabilitation stays whenever a rehabilitation patient returns to the
same IRF within a short period of time. Cases that have a one-day
interruption appear to differ from other interrupted cases. Only 29
percent of patients whose program is interrupted for less than a
calendar day are discharged. Only 5 percent of discharges go to acute
care. Paying for both discharges in a very short interruption raises
issues of fairness, and it seems very reasonable to bundle these cases.
They would be slightly overpaid if each part of the stay received a full
discharge payment. There are very few such cases, so bundling is
unlikely to pose a substantial problem to any hospital.

Beyond the same-day interruptions, there is no analytical reason to
choose any one particular period as the bundling criterion. We find
discharges for over 80 percent of all recorded interruptions up to 30
days in length. For any period up to 10 days, we would pay about the
same 70 percent of costs as we do for the interval proposed in the NPRM
(three calendar days, including the discharge day). Using a per
discharge payment, including a per diem payment for transfer cases,
accurately matches payments to cost.

It is important to consider how incentives would change with a
bundling policy. The substantial majority of patients whose interruption
extends beyond a calendar day are admitted to acute care. The causes of
these interruptions are largely cardiac and other problems prevalent in
this sick population and are only rarely surgical problems. The
incentive for these patients is to not readmit them to rehabilitation
until the bundling period ends. If patients remain in acute care for

another day, little is probably lost from their rehabilitation program. The incentive may therefore be strong--particularly for units, because the hospital will gain much more than the cost of the extra day of acute care. So the only thing that is lost is some wasted resources. In the worst case, the patient might instead be transferred elsewhere and might not return to rehabilitation. Bundling the acute care costs would only aggravate the problem.

In general, we believe that matching payments to expected costs as closely as possible is the best way to reduce opportunities for profitable gaming and to reduce the effect of financial considerations in clinical decisions. We therefore recommended that discharges for an interruption of more than 24 hours not be bundled.

There are other possibilities for bundling. One is for the small number of cases, less than one-half of one percent, that are transferred from one rehab facility to another. One might consider paying a single payment to cover both facilities and letting the hospitals work out the split. This is actually the transfer policy that was proposed in FY 1983 for the acute care PPS. It raised serious objections in the field due to the administrative implications, and we expect it would raise similar objections here. The expert panel on our previous study said that transfers from one rehabilitation facility to another were almost always made at the request of the patient. This increases the administrative problem of the facility under such a policy, because the sending hospital has no control over, and likely no financial arrangement with, the receiving hospital. The only concern here would be if some pairs of hospitals were to look on this policy as an opportunity to game the system. Thus we recommended active monitoring of the source and destination of such transfers, with action taken only if abuse is detected. In the meantime, the same transfer policy should be used for transfers to a rehabilitation facility as for transfers to acute care hospitals, LTC hospitals, nursing homes, and SNFs.

Bundling of payment across all post-acute providers and/or with the acute stay provider is certainly a desirable long-term goal. We believe that such a policy should be considered over a longer time frame rather than implemented in the near future. Authority to change the payment for

acute care and other post-acute providers because of the patients' use
of a rehabilitation facility may not be supported by current
legislation.

6. RELATIVE CASE WEIGHTS

For any particular hospital, the payment for each case will be proportional to the relative weight assigned to the patient's CMG. To ensure that beneficiaries in all CMGs will have access to care and to encourage efficiency, we want to calculate weights that are proportional to the resources needed by a typical case in the CMG. So, for example, cases in a CMG with a weight of 2 will typically cost twice as much as cases in a CMG with a weight of 1.

The average of the relative weights for a set of cases is called the case mix index, or CMI. The CMI adjusts hospital payment for case mix. The CMIs for cases at different hospitals can be compared to describe the relative costliness of each hospital's case mix.

In this section, we describe the method chosen for weight calculations and related analyses. Since relative weights and CMIs depend on the underlying case classification system, we report CMS's decisions about case classification in the first subsection below. Because the FRGC methodology for typical cases has been modified[1] and the system includes classification of atypical cases, we call the resulting classification system *case mix groups* (CMGs).

Because weights constructed with our data show CMI compression, we define compression and discuss its causes in the second subsection. The next subsection discusses policy options for calculating relative weights for the CMGs and reviews the findings from our interim report about these options. We then describe the data and methods that we used to calculate the weights for the final rule. The subsection on results presents the final comorbidity effects, describes the extent of compression in the data, and gives the evidence for one cause of compression. We end by presenting a remedy for CMI compression and summarizing the implications of our research.

---

[1] The motor score has been redefined; the rule for accepting CART splits has been changed; comorbidity tiers are now used; and long-stay transfers are included in the groups.

CASE CLASSIFICATION DECISIONS

Typical Cases

We defined *typical cases* as non-transfer cases that are discharged alive and that spend more than three days in the rehabilitation facility. These cases were classified using a variation of the FRGC methodology as discussed in Sections 3 and 4 (in particular, see Tables 3.1, 3.16, and 4.10).

Atypical Cases

Other rules for case classification are as follows:

- All non-transfer cases that stay in the hospital for three days or less, including death cases, are placed in a single group.

- Death cases that stay for four days or more are placed in one of four groups based on LOS and whether or not they are orthopedic cases. *Orthopedic cases* are defined as those with a primary impairment that causes assignment to one of the orthopedic RICs (7,8, or 9). Each of the sets of orthopedic and non-orthopedic death cases are then subdivided based on whether they stay less than the mean LOS for that set of death cases.

- *Short-stay transfers* are defined as cases discharged to a hospital or to an institution paid as a nursing home by either Medicare or Medicaid and that stay no longer than a half day less than the mean LOS of cases assigned to the same FRGC and paid as typical cases. Short-stay transfers are paid a per diem equal to the average daily payment for a typical case in the same FRGC plus an additional case-level payment of one-half of the per diem.

- *Long-stay transfers* are classified just as typical cases are.

Bundling

CMS decided that discharges where the patient returns to the same rehabilitation facility within three days[2] should be bundled with the subsequent discharge for purposes of case classification and payment.

---

[2] The day of discharge is counted as day 1, so this is the same as return on the day of discharge or either of the following two calendar days.

To develop relative weights consistent with this decision, we proceeded as follows. We began by identifying all MEDPAR discharges for all hospitals with at least one discharge in our matched FIM-MEDPAR data set (including both matched and unmatched discharges). We then matched up every set of MEDPAR records where the same person returned to the same rehabilitation facility within three days.

The rules that classify interrupted stays are the same as those we used to analyze bundling in Section 5. Briefly: The impairment code and FIM item scores come from the FIM record that matched the first discharge or the FIM record that matched the second discharge when it recorded an admission within one day of the first. We used the last discharge in the interrupted stay to determine discharge status for the bundle. The comorbidity tier assignment was based on all secondary diagnoses that appeared on any matched discharge in the stay. Finally, we summed days and cost across each discharge to assign the LOS and cost for the interrupted stays.

The 3,021 bundles in our 1998-1999 sample came from 6,152 discharges. The ratio of discharges to bundles in our analysis file is 1.0066, quite similar to the number for 1996-1997 (1.0062). The ratio of all discharges to all bundles at the sample hospitals is slightly higher--1.0097.

COMPRESSION

The term *weight compression* is used to indicate that a set of weights has the property that the weights of high-weighted cases are too low--i.e., high-weighted cases are undervalued relative to resource needs, and low-weighted cases are overvalued relative to resource needs. Similarly, CMI compression indicates that hospitals with high CMIs have costs that are higher relative to their CMI than do hospitals with lower CMIs.

Weight compression is one of the possible causes of CMI compression, but not the only one. Hospitals with high CMIs could, on average, spend more resources on low-weight cases than do hospitals with a low CMI. If so, the relative weights could be exactly proportional to resource use, but the weights would still exhibit CMI compression.

To measure CMI compression, it is usual to regress the log of average cost at a hospital on the log of the CMI. The coefficient on the CMI shows how much the cost increases with increasing CMI. If the weights are neither compressed nor decompressed, the coefficient will be 1. A value greater than 1 would indicate CMI compression--hospitals with a high CMI (i.e., patients who need more than average amounts of resources) have higher costs relative to their CMI than do hospitals with lower CMIs. For example, if the coefficient is 1.2, a hospital with a CMI that is 10 percent higher than another will on average cost 12 percent more.

If one assumes that there is no correlation between hospital case mix and either efficiency or output, a finding of CMI compression can allow one to infer weight compression. It seems unlikely that hospitals with a high CMI are more inefficient on average than those with a lower CMI. Further, we found in the interim report that increasing the number of classification categories had only a tiny effect on CMI compression. Consequently, we do not think that the hospitals with a high CMI are consistently getting patients who are more costly within each CMG and who are identifiable from their FIM scores, comorbidities, and/or age.

Weight compression is a more likely explanation for CMI compression than is inefficiency or case selection. Three reasons for such weight compression are usually given:[3]

- Each hospital is assigned just one per diem for routine cost, yet per diem nursing costs may vary by CMG.
- A single cost-to-charge ratio is used within each ancillary department. Some observers believe that acute care prices are set so that low-cost services subsidize higher-cost services, even within a particular ancillary department. To the extent that this is true in rehabilitation facilities, the cost-based weights of CMGs that use those low-cost services (which might be low-cost CMGs) will be overestimated, and the cost-based weights of CMGs that use the higher-cost services will be underestimated.
- Errors in classification of cases into CMGs will tend to make the weights more similar than they should be. We believe that in our

---

[3] See Lave (1985) or Carter and Rogowski (1992).

data set, comorbidities may be undercoded and FIM item scores may
be in error, thus causing classification errors.

We believe that a fourth factor is also at work in our data: the
bundling of ancillary services into the routine per diem. We will
present data consistent with this hypothesis in the Results subsection.

REVIEW OF PREVIOUS ANALYSES

Relative weights could be calculated in a variety of ways. In our
interim report (Carter et al., 2000, Section 5), we examined
alternatives that differed in the measure of resource use (charge versus
cost), the method used to control for variations in costs across
facilities (hospital-specific standardization versus payment factors),
and whether the weight algorithm should account for the proportion of
outlier payments in the payment group. We review these results and other
relevant literature here.

Resource Measure

Cost-based weights were used in the initial acute care PPS
implementation, but charge-based weights have been used since FY 1986.
Charge-based weights can be updated more quickly than cost-based weights
because they do not require the filing and auditing of cost reports, a
process that may take up to two years. Charge-based weights can be
updated based on discharge data that are typically available after only
six months.

Cotterill, Bobula, and Connerton (1986) found charge- and cost-
based weights for the acute care PPS to be quite similar. In addition,
some exploratory analyses found that charge-based weights better
estimated cost for hospitals with a high CMI, although charge-based
weights were still compressed (Thorpe, Cretin, and Keeler, 1988). By the
fifth acute care PPS year, CMI compression in charge weights had
disappeared (Carter and Rogowski, 1992) and those authors predicted that
charge weights would become "decompressed" in the future.

The distribution of patients with comorbidities varies across
rehabilitation facilities. Thus our criteria of access and provider
fairness required that we measure additional costs associated with
comorbidities as accurately as possible. However, our interim report

found two pieces of evidence to indicate that charges measure the relative costs of comorbidities less accurately than do accounting costs. First, cost-to-charge ratios vary substantially across the major departments in ways that should affect measurement of the relative cost of cases with comorbidities. Second, t-statistics measuring the effect of comorbidities on standardized charges within each RIC were noticeably lower than those for standardized accounting costs.

In our interim report (Table 5.5), we also found that charge-based standardized weights showed just as much CMI compression as cost-based, standardized weights did. We therefore recommended that cost-based weights be used for the IRF PPS, and CMS followed our recommendation in both the NPRM and the final rule.

Controlling for Hospital Costs

Both charges and costs are affected by hospital characteristics for which the IRF PPS could adjust payment--i.e., the factors that go into the calculation of the facility payment factor. In the standard method used to calculate DRG weights, charges are first standardized by the hospital's payment factor.

The payment factors, however, capture only a small part of the variation across hospitals in costs for any specific DRG. A method that accounts for more of the cross-hospital variation in costs than the standard method is the hospital specific relative value (HSRV) method. The HSRV method differs from the standard method in that a hospital's costs are not standardized using its payment factor, but instead are standardized using hospital-specific costs and the hospital's case mix index. Hospital identity is entered into the regression that estimates comorbidity effects. The HSRV method should be superior to the standard method because it produces weights that reflect relative accounting costs. However, because within-department pricing rules vary and are unknown, there is no guarantee that, after averaging across hospitals, the HSRV method will yield relative weights closer to relative actual costs (rather than accounting costs).

One problem with the standard method of calculating weights is that one must know the weights in order to calculate the payment factor and

one must know the payment factor to calculate the weights. The HSRV
method can be used to get initial weights, which could be used to obtain
payment factors. These, in turn, could be used to obtain standard
weights and then a refined payment factor.

In our study of the model PPS (Carter et al., 1997), we evaluated
both methods of controlling for hospital costs. Using the standard
method and the facility payment factor to standardize charges provides
similarity to the acute care PPS. However, Hosek et al. (1986) found
substantial correlation in charges and length of stay for patients
treated at the same facility even after controlling for impairment
group, functional status, and facility characteristics. Further, the
TEFRA rules have allowed newly certified units to recover larger costs.
Thus, on a priori grounds, one might prefer the HSRV method in this
circumstance.

In our interim report, we compared HSRV cost weights to
standardized cost weights. We found that controlling for hospital
identity produced larger t-statistics and smaller standard errors for
the comorbidity effects than using payment factors as a hospital
control. We inferred that hospitals systematically vary in their
costliness in ways beyond those accounted for in the payment adjustment
and that the HSRV method results in more-precise estimates of the
effects of comorbidities on costs than does the standardization method.

Differences between the HSRV weights and the standardized cost
weights were small but occurred throughout the range of weights. In
particular, there was no noticeable correlation between the magnitude of
the weight and the difference between the HSRV weight and the
standardized cost weight. Although the regression coefficient that
measures the CMI compression in the standardized cost weight is lower
than that of the HSRV weight, it was quite significant.

In the interim report, we recommended that HCFA use the HSRV
method, and our recommendation was accepted. For the final rule, we
recommended that the HSRV method be the basis for the weights, but that,
as we will discuss later in this section, the weights be adjusted to
eliminate CMI compression. This recommendation was also accepted.

Fair Weights

The term *fair weights* has been given to the result of weight calculations that adjust for the additional reimbursement received by outlier cases. Under the acute care PPS, DRG weights are calculated to be proportional to the resources used by the average case in the DRG. Fair weights could be calculated so that, instead, total PPS standardized reimbursement (including outlier payment) would be proportional to the resources used by the average case in the DRG. Such fair weights have been considered by ProPAC (1994) and are sometimes known as "DRG Specific Outlier Funding," or DSOF.

Outlier payments will be a higher percentage of IRF PPS payments in some CMGs than in others. This is appropriate insofar as it provides more outlier payment for the cases that would otherwise cause the highest losses. In our interim report, we showed that outlier payments improve the accuracy of payment at the case and hospital level. Implementing fair weights would worsen compression and lower accuracy. Thus we believe that we should not use fair weights in the initial system. We did, however, take into account outlier payments at each hospital when we determined the factor used to eliminate compression.

CALCULATING HSRV WEIGHTS FOR THE IRF PPS

The first step in the calculation of CMG weights was to estimate the effect of comorbidities. The second step was to adjust the costs of each discharge for these effects. We then used these adjusted costs to calculate "relative adjusted weights" in each CMG by the HSRV method. The final steps were to calculate the weight by using the effects of comorbidity to modify the relative adjusted weight and then to normalize the weights to 1.

Data

The data we used to estimate comorbidity effects differ from those used in Section 4 because for this calculation we grouped long-stay transfers with typical cases in the same CMG. Long-stay cases will thus affect the relative weight for each tier. The data also reflect the bundling of discharges into cases as discussed above, although this causes a much smaller difference than the one due to the inclusion of long-stay transfers.

To calculate weights, we used all cases in our analysis set, including all transfers, atypically short stays, and in-hospital deaths.

We report comorbidity effects only for the combined 1998-1999 data that were used for the weights that appear in the final rule. Because of the low incidence of comorbidities, we used two years of data to calculate the comorbidity effects. Cases outside three standard deviations of the mean of log (costs) in the CMG were excluded from the regressions.

We calculated separate weights for 1998 and 1999 using the same comorbidity effects. The weights in the final rule were based on the 1999 data. Statistical outliers were removed, as described below.

Estimating the Effect of Comorbidity

As we saw in Section 4, the presence of each comorbidity tier multiplies a case's expected resource use by the same amount for each CMG in the same RIC. Thus, we used the log transform and regression to calculate the comorbidity effect for each RIC.

We performed a single regression with the case as the unit of analysis. The dependent variable was the log of case cost.[4] The independent variables were indicator variables for each CMG (except one, since we have an intercept term), indicators for each provider except one, and a three-element indicator vector for each RIC with each element describing whether the case is in one of the three tiers. The comparison groups for the comorbidity variables consist of cases in the same CMG and hospital with no relevant comorbidity.

If the three-element transposed vector of coefficients for the RIC containing CMG k is a_k, then the weight for a case in CMG k is given by

$$W(k,x) = \exp(a_k x) * w_k, \qquad (6.1)$$

where

w_k = the comorbidity adjusted weight for the CMG,

x = a three-element vector of indicator variables describing the
    case's comorbidity status (i.e., x has one element for each

---

[4] Because we used the HSRV method, which adjusts for hospital identity, adjusting for wages would not further affect the comorbidity coefficients.

tier which is 1 if the patient is in that tier, and 0 otherwise).

Adjustment for Comorbidity

The second step in the calculation of weights was to adjust the resource use for each case to eliminate the effect of comorbidities. The adjusted resource use for a discharge, with vector comorbidity x is

$$A = cost/exp(a_k x), \tag{6.2}$$

where the coefficient vector a_k comes from the comorbidity regression.

We then used the adjusted resource-use values for each case to calculate the relative adjusted weight in each CMG k, w_k, by the HSRV method. The adjusted resource use for short-stay cases and in-hospital deaths is the same as their actual resource use.

Relative Adjusted Weights

In the HSRV method, the adjusted costs are standardized at the hospital level using hospital-specific costs--so costs for a patient at a hospital with high average costs for its patient mix are counted as less resource use than costs at a hospital with low average costs for its patient mix. The average weight, or CMI, is used to account for the variation in patients across hospitals.

In counting the number of cases at the hospital and calculating the hospital's CMI, we counted short-stay transfer cases as a fraction of a case. This fraction was set to equal the ratio of the LOS of the transfer case plus 0.5 day to the average LOS for typical cases and long-stay transfers in the same CMG tier combination. This is the same method used by HCFA in the acute care PPS.

The total adjusted cost for each case was divided by the average adjusted cost for the hospital in which the case occurred. The resulting ratio was then multiplied by the hospital's CMI to produce a hospital-specific relative value. This relative value can be viewed as a cost that has been standardized by the hospital's own costliness, in contrast to the standard method in which the cost estimate is standardized by the payment factor.

The process of calculating the weights is iterative. We first chose initial values for the CMI of each hospital. CMG-adjusted weights were then set in proportion to the average value of the HSRV, resulting in a new CMI for each hospital and therefore new HSRVs. We continued the process until there was convergence between the weights produced at adjacent steps, for instance, when the maximum difference was less than 0.0001. Earlier work had shown that the algorithm is not sensitive to starting values of the CMI (Rogowski and Byrne, 1990; Carter and Rogowski, 1992).

After the first iteration, we eliminated statistical outliers-- defined as cases with HSRV standardized costs that differ from the CMG mean by more than three standard deviations in the log scale.

Relative Weights for CMG-Comorbidity Combinations

The next step in the algorithm was to calculate a relative weight for each relevant combination of CMG and comorbidity tier, using Eq. (6.1). The final step of the algorithm was to multiply by the normalizing constant so that the average weight per case is 1.

RESULTS

Effect of Comorbidity

Table 6.1 shows the comorbidity effects in each RIC in the regressions that were used to calculate relative weights. For each CMG, the line for the RIC containing that CMG gives the vector a_k used in Eqs. (6.1) and (6.2).[5] The R-squared is much larger than that from Model 4 in Table 4.7. Most of the increase is due to the inclusion of controls for provider, as is required by the HSRV weight method.

---

[5] For atypical short stays and in-hospital deaths, the vector a is uniformly 0 (so exp(ax) = 1 for any x).

Table 6.1

Effect of Comorbidity Tiers in Weight Regression

RIC	Tier 1		Tier 2		Tier 3		F-test for Significance of Comorbidity Vector	
	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	F-value	Sig-nifi-cance
1	0.2136	11.09	0.1034	10.87	0.0551	11.22	114.28	<.0001
2	0.2200	4.89	0.1648	8.42	0.0859	4.99	32.52	<.0001
3	0.2940	7.34	0.1354	9.25	0.0928	8.10	53.53	<.0001
4	0.2954	4.83	0.2325	6.95	0.1741	5.83	30.18	<.0001
5	0.3473	6.80	0.2627	17.11	0.1499	11.29	140.43	<.0001
6	0.3908	9.77	0.1262	11.96	0.1052	11.30	106.72	<.0001
7	0.1630	4.07	0.1617	15.48	0.1186	15.02	150.63	<.0001
8	0.2328	5.73	0.1885	23.10	0.1498	21.55	327.25	<.0001
9	0.2931	4.94	0.2036	13.24	0.1447	13.08	115.09	<.0001
10	0.1771	3.30	0.1816	13.70	0.0921	11.27	88.47	<.0001
11	0.6436	2.60	0.1473	4.20	0.0752	3.02	8.69	<.0001
12	0.4514	4.82	0.1666	7.51	0.1045	6.63	38.56	<.0001
13	0.3270	2.96	0.1585	5.69	0.1449	6.65	26.2	<.0001
14	0.3325	10.38	0.2212	19.17	0.1041	11.01	171.45	<.0001
15	0.1941	2.86	0.1431	9.18	0.0460	2.77	31.59	<.0001
16	0.2765[b]	2.50	0.2319	8.92	0.1776	8.96	51.38	<.0001
17	0.3275	5.42	0.2250	6.80	0.1218	4.84	30.09	<.0001
18	0.0877[a]	0.85	0.1953	3.23	0.0883[a]	1.48	3.93	0.008
19	0.2664	2.78	0.1194[b]	2.36	0.0971[a]	1.88	4.64	0.003
20	0.3282	19.78	0.1768	27.82	0.1063	20.70	430.87	<.0001
21	0.2110[a]	1.17	0.1167[a]	1.37	-0.1759[b]	-2.32	3.46	0.016

[a] $p > 0.05$.
[b] $0.05 > p > 0.01$.
Note: Regression used 394,648 cases from 1998 and 1999. R-squared is 0.5232. Regression also controlled for CMG, provider, and year. F-statistics have 3 degrees of freedom and come from separate regressions within each RIC that do not control for provider number.

The comorbidity vector is statistically significant in every single RIC, usually with $p < 0.0001$. As shown by the t-statistics, the effect of each tier is measured with reasonable precision in almost all RICs. In almost every RIC, the tier coefficients are ordered correctly--i.e., tier 1 coefficient > tier 2 coefficient > tier 3 coefficient > 0.

The exceptions to monotonicity require attention. First, in RICs 10 and 18, the tier 1 effect is smaller than the tier 2 effect but also statistically indistinguishable from the tier 2 effect. Thus we averaged the two coefficients, weighting by the number of cases in each tier, and

used the average as the elements of a_k for both tier 1 and tier 2. Second, the burn RIC (21) is too small to measure tier effects accurately. (There are only six cases in tier 1 and 31 cases in tier 2.) If we pool the two tiers, we get a coefficient of 0.16 with a t of 1.46 ($p < 0.15$ in a two-tailed test; $p = 0.07$ in a one-tailed test). We believe that the coefficients on tier 1 and tier 2 are more accurate than 0, so we used the tier 1 and tier 2 effects. The tier 3 effect was measured as negative and set to zero, so the weight for tier 3 cases in RIC 21 is the same as the weight for a case in the same CMG with no relevant comorbidity.

For completeness, we also checked that there was an interaction of each tier and RIC within this data set. Such is clearly the case. The three interactions of RIC with each of the tiers were all significant with $p < 0.0001$.

Compression Remains in the Unadjusted HSRV Weights

The weights are intended to adjust payments for differences in case mix across hospitals. If one assumes that there is little correlation between average case mix and hospital inefficiency, then one would want average hospital costs to be directly proportional to the CMI at the hospital. We regressed the log of wage-adjusted costs on the log of the CMI and wanted the coefficient on the CMI to be 1. A value higher than 1 would indicate CMI compression--hospitals with a high CMI (i.e., with patients who need more than average amounts of resources) have higher costs relative to their CMI than do hospitals with lower CMIs. The regressions control for rural location and the percentage of low-income patients at the hospital, using the form of the variables in the final payment adjustment (see Section 7). We weight each hospital in proportion to the number of sample cases at the hospital, with short-stay transfers counted as a fraction of a case.

Table 6.2 shows that increases in the weights underestimate the magnitude of increases in costs in all years. CMI compression continues to exist in the weights in 1998 and 1999 data. To measure pure trends, the first three rows show results using the FRG and comorbidity

Table 6.2

**Compression in Hospital Case Mix: Coefficient on Log CMI in Regression
of Log Hospital Wage-Adjusted Cost per Case, by Data Set, Year,
and FRGC Definition**

Cases Included in Hospital Average Cost and CMI	FRGC Definitions	Year	Coef.	Std. Error	T-statistic for Hypothesis That Coef. Is 1	R-Squared
All	Interim report	1997	1.2320	0.0723	3.21	0.4091
All	Interim report	1998	1.2202	0.0769	2.86	0.3578
All	Interim report	1999	1.2509	0.0795	3.16	0.3500
All	Final recomm.	1998	1.2258	0.0719	3.14	0.3844
All	Final recomm.	1999	1.2305	0.0733	3.14	0.3730
Short-stay transfers excluded	Interim report	1997	1.2669	0.0766	3.48	0.3946
Short-stay transfers excluded	Interim report	1998	1.2363	0.0808	2.93	0.3456
Short-stay transfers excluded	Interim report	1999	1.2401	0.0827	2.90	0.3332
Short-stay transfers excluded	Final recomm.	1998	1.2322	0.0755	3.08	0.3684
Short-stay transfers excluded	Final recomm.	1999	1.2177	0.0763	2.85	0.3545

Note: Based on 618 hospitals in 1997, 667 in 1998, and 694 in 1999.
Average cost and case mix excluded short-stay transfers. Regressions
weighted by number of cases at hospital. All t-statistics significant at
$p < 0.01$. Regression also included an indicator for rural location and
hospital's low-income percentage.

definitions from the NPRM.[6] There is only a small difference among the
three years. The next two rows show the effect of our new FRGC
definitions and transfer payment rule. There is a noticeable improvement
in R-squared, but little change in compression. Under our best-case
classification system, the coefficients on the log of the CMI were
1.2258 in 1998 and 1.2305 in 1999. The t-statistic in the table shows

---

[6] The weights are the HSRV-V4 weights from the interim report,
which is the same method used with the new CMGs.

that both of these numbers are significantly above 1 (p < 0.01). Thus a hospital with a CMI that is 10 percent higher than the CMI of another hospital will have costs that are typically 12.4 percent higher.

In the last half of the table, short-stay transfer cases have been eliminated to demonstrate that our analysis was not confounded by their weights, which are slightly lower than would be expected from their costs. The conclusions are roughly similar. The R-squareds are lower when we exclude transfer cases, showing that the transfer rule helps to explain variance across hospitals in cost per case.

Bundling of Ancillary Services Could Cause Weight Compression

It is likely that problems in measuring cost per case in each FRGC are a major factor in compressing weights, which in turn leads to CMI compression. In this subsection we examine data related to one of the factors that could cause weight compression--the bundling of ancillary services with routine care. We have no data with which to examine whether cost-to-charge ratios vary for different services within an individual ancillary department of an individual hospital but believe that such effects, if any, are small. We also have no data with which to examine the effect of variation across patient types in nursing time on the measurement of costs. Because comorbidity data were not used, we expect that they are undercoded in our data, but we have no direct test.

We next present data that we believe strongly suggest that a substantial fraction of hospitals include many ancillary services in routine costs on their Medicare cost report. Further, the data are consistent with this bundling being a major cause of weight compression.

The bundling of ancillary services into routine costs, in itself, would not be a problem if these services were actually delivered to every patient on a daily, or frequent and periodic, basis. For example, in some hospitals linen charges are counted under "supplies," and in others, under routine cost. This is not a problem since almost all patients have the same daily linen costs and thus the bundling does not cause large errors in estimating the cost of each patient. The problem arises when the services are delivered only to some patients and/or are delivered only rarely or not regularly (so that an increase in LOS does not cause a corresponding increase in the use of the services). In this

case, the cost of patients who are receiving these services is underestimated and the cost of other patients is overestimated.

We do not have access to details of the charging structure of the sample IRFs, so we cannot directly measure either the extent to which they are bundling patient-varying ancillaries into the routine per diem or the effect of such bundling on weight compression. Instead, we ask a series of questions. First, is the use of ancillary costs strongly and positively related to case weight? If so, and if bundling of patient-varying ancillary services into routine costs occurs, it would cause weight compression. Second, do hospitals vary in the percentage of costs attributable to ancillary services, even after controlling for case mix? Third, do hospitals with lower-than-expected ancillary costs have higher-than-expected routine costs? We believe that if the answers to the second and third questions are positive, it is reasonable to infer that the hospitals are bundling ancillary services into routine case costs. Finally, we look at the within-hospital variation in costs as a function of the percentage of ancillary services.

Ancillary Costs Are Correlated with Case Weight

Per diem costs for a case are the costs associated with adult and pediatric routine care or special care. In inpatient rehabilitation, almost all costs are routine care; special-care units are used very rarely. *Ancillary costs* are those associated with all other departments, such as therapy, pharmacy, or radiology. We calculated the two kinds of costs separately and summed them to get total cost for the case.

We averaged ancillary costs and per diem costs within each FRGC and then found the fraction of costs that were ancillary. In our sample, 41.4 percent of all Medicare rehabilitation costs arose from ancillary services and 58.6 from routine costs. As expected, cases with comorbidity used a higher proportion of ancillary services. For comorbidity tiers 1 through 3, the proportion of costs from ancillary services was 47.7, 44.1, and 42.9, respectively, whereas cases with no relevant comorbidities had 40.5 percent of costs from ancillary services. Within each RIC and comorbidity tier, the percentage of costs from ancillary services increased with FRGC weight. For example, for cases with no comorbidities in RIC 1, the fraction of costs for

ancillary services was only 38.8 percent in the lowest-weighted FRG, whereas it was 43.4 percent in the highest-weighted FRG.

We used the case-weighted correlation between the fraction of costs in each FRGC that are ancillary and the FRGC weight to summarize the overall relationship between case weight and percent ancillary costs. This correlation is substantial: 0.61 in 1998 and 0.63 in 1999 ($p = 0.0001$ each year).

Hospital-Level Variation in Ancillary Costs

We next examined differences among hospitals in the extent to which ancillary costs are bundled into routine costs. For each CMG, we first calculated the average fraction of costs that come from ancillary services. Then, for each case we calculated the expected ancillary cost if the hospital used ancillary services in exactly the same proportion as the average for that CMG. Then we summarized these expected costs at each hospital and compared it to the actual ancillary cost per day at the hospital. The ratio of these two costs is a measure of the extent to which the hospital's ancillary costs are typical.

We calculated the average ancillary cost per day at hospitals in each of the four quartiles of the ratio of ancillary costs to expected ancillary cost. In Table 6.3, we compare daily ancillary cost with that expected from the CMG distribution of the hospitals and with the routine cost per diem. The third column of Table 6.3 shows the average daily ancillary cost at the hospitals in each quartile and year. Ancillary costs increase by about 50 percent from the first to the fourth quartiles each year. The fourth column shows the ancillary cost that would have been expected given the hospital's CMG mix and its own measured case costs. In 1999, hospitals in quartile 1 had ancillary costs about $100 per day (= 364 - 262) less than expected, whereas hospitals in quartile 4 had ancillary costs of $75 per day more than expected. A similar pattern is found in both years.

The lowered daily ancillary cost in quartile 1 is more than offset by higher per diem costs, and the higher ancillary cost in quartile 4 is offset by lower per diem costs. These data are evidence of hospitals using different rules to categorize a substantial volume of services into routine and ancillary.

Table 6.3

Daily Costs by Year and Quartile of Percentage of Hospital Costs That Are Ancillary

Year	Ancillary Percentage Quartile	Number of Cases	Average Daily Ancillary Cost	Expected Daily Ancillary Cost	Average Per Diem Cost
1998	1	45,179	246	330	568
1998	2	58,558	272	288	439
1998	3	64,527	289	276	389
1998	4	57,937	365	291	350
1998	All	226,201	295	294	428
1999	1	47,797	262	364	619
1999	2	62,352	279	296	439
1999	3	64,767	304	290	397
1999	4	72,545	379	304	356
1999	All	247,461	312	310	438

Note: Expected daily ancillary cost is based on CMG distribution at the hospital.

The remaining question is whether the services being classified differently are provided daily (or frequently and periodically) to all patients. Here, we have only indirect evidence that they are not. As shown in the third column of Table 6.4, in each year, hospitals in the

Table 6.4

Within-Provider Standard Deviations of Log-Cost and Log-Weight and Regression Mean Standard Error

Year	Ancillary Percentage Quartile	Within-Provider Standard Deviations Log-Cost	Log-Weight	RMSE from Regression of Log-Cost on Log-Weight and Provider
1998	1	0.6034	0.4812	0.4068
	2	0.6174	0.4857	0.4133
	3	0.6187	0.4748	0.4175
	4	0.6085	0.4786	0.4159
1999	1	0.5978	0.4798	0.4006
	2	0.6232	0.4831	0.4138
	3	0.6095	0.4736	0.4160
	4	0.6154	0.4786	0.4231

first ancillary quartile have less variation across patients in estimated costs than hospitals in any of the other quartiles. If we control both for provider and case weight, the remaining variation in patient costs (RMSE) is noticeably lower in quartile 1 than in other quartiles.

Bundling Patterns Probably Contribute to Weight Compression

We have shown that hospitals vary in the percentage of costs attributable to ancillary services, even after controlling for case mix. Further, hospitals with lower-than-expected ancillary costs have higher-than-expected routine costs. We believe that it is reasonable to infer that some hospitals are bundling ancillary services into routine case costs. Further, the hospitals with the lowest fraction of ancillary services have lower-than-average variation in patient costs--thus it is likely that some of the bundled ancillary services actually vary across patients and patient days.

Ancillary costs are strongly and positively related to case weight. Therefore the bundling of patient-varying ancillary services into routine costs results in an underestimate of the cost of high-weight cases and an overestimate of the cost of low-weight cases, thus causing weight compression.

POTENTIAL REMEDY FOR COMPRESSION

We have seen that our weights have CMI compression. At this point we are left with several likely inferences about the cause of this CMI compression:

- The CMI compression is likely due, at least in part, to weight compression.
- The weight compression is likely caused by errors in measuring the cost of individual cases at specific hospitals, possibly aggravated by undercoding of comorbidities and miscoding of FIM scores.
- The measurement of hospital average cost per case is likely more accurate than the measurement of average case cost.
- Based on the analysis in the interim report, charge-based weights and cost-based weights are also compressed. Indeed, insofar as the causes of the compression are variations in nursing costs, coding errors, or

the bundling of ancillary costs into routine costs, there is no reason for them to be less compressed.

Although the measurement of individual case costs is problematic, weights derived from individual case costs are in fact reasonable predictors of hospital-level costs. It is possible to directly address the CMI compression problem by spreading out the weights. Before proceeding with this approach, we examined the pattern of the increase in log costs with log CMI. For the years 1998 and 1999, we regressed the log of costs, standardized for wages and CMI, on the facility adjustment variables of low-income percentage and rural hospital. Then we plotted the residuals against the log of the CMI. The full scatterplots are very crowded and therefore somewhat conceal the relationship. So we ordered the hospitals by CMI and then grouped them into 30 groups with similar values of CMI and approximately equal proportions of cases. Then we averaged the residuals and plotted them, as shown in Figure 6.1. Although a large number of hospitals with CMIs near 1 (log(CMI) near 0) have variable costs, the linear pattern is clear in both years' data.

Linearity in the log implies a multiplicative effect in the untransformed CMI. It would be possible to address compression either with a facility-level adjustment or with a case-level adjustment. However, the evidence suggests that the weights themselves are compressed, so an adjustment to the weights would be preferable to an adjustment to the CMI. It also would be easier to implement.

To determine the adjustment, we used the simulation that accounts for the exact formula used for the facility adjustment and for outlier payments. Table 6.5 shows a regression of the log of average costs in 1999 at each hospital on the log of average payment at the hospital. The coefficient of 1.25 is close to the 1999 result of 1.23 for all cases in Table 6.2. However, this does not account for the fact that both costs and payment are higher at hospitals that have many outliers. Consequently, the second row of this table shows a regression of the log(average cost minus average outlier payment) on log(average payment minus average outlier payment). This regression shows the relation between costs not compensated by outlier payments and the weights. The R-squared is much lower because the predictive ability of outlier payments has been removed.
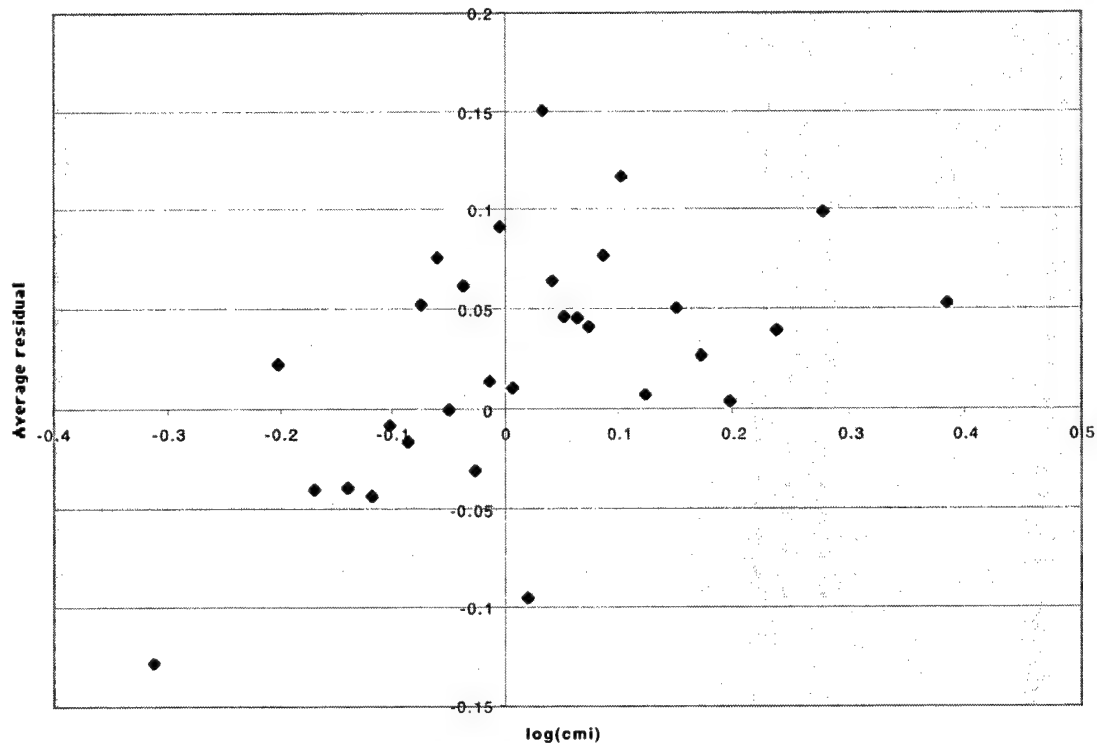
Figure 6.1--Residuals of Regression of Log Wage and CMI Adjusted Cost on Payment Adjustment Factors (1999 data)

Table 6.5

Regression of Log Hospital Cost per Case on Log Payment per Case (1999 data)

	Intercept	Coefficient	t-stat.	R-squared
All data	-2.388	1.2532	30.47	0.5729
Minus outlier payments	-0.960	1.1007	23.35	0.4407

Note: Regressions are weighted by number of cases at the hospital, with short-stay transfers counted as a fraction of a case. Facility adjustment is from Model 10 in Table 7.12.

As one can see from Table 6.5, a 1 percent increase in the CMI results in a 1.10 percent increase in non-outlier costs in 1999. So, if one wanted to have average hospital payments that were proportional to average hospital cost per case, one could change the weights so that

$$nw(i) = w(i) + 0.10(w(i)-1) \qquad (6.3)$$
$$= 1.10\ w(i) - 0.10.$$

This formula increases all weights that are greater than 1 and decreases all weights that are less than 1 by stretching the difference between the weight and the average value of 1. The effects on weights are small but become larger as the weights get further from 1. For example, a CMG with a compressed weight of 2 would have a "decompressed weight" of 2.1. The effect on hospital CMIs is even smaller. The 75th percentile of CMI is 1.11. Such a hospital would see its CMI increase to 1.122. The 95th percentile hospital would see its CMI increase from 1.315 to 1.3465 (a 2.4 percent increase).

Most resulting weights appear plausible. However, an exception must be made for atypical short-stay cases, whose payment would drop by about half--way below any reasonable estimate of cost. We limited the change so that no CMG would have more than a 15 percent decrease in weight. The atypical short-stay cases were the only CMG affected.

The advantage of this weight system is that it results in average payments per hospital that vary directly with the hospital's average costs. The system explains hospital-level cost as well as the initial set of weights did. (In 1999, the new CMI produced an R-squared of 0.3722 rather than 0.3730.) The disadvantage is that we do not know how it affects the accuracy of the case-level payments. However, one who believes that weight compression is largely due to the bundling of ancillary services into routine costs, or that nursing intensity is higher for cases with higher weight, can infer that the decompressed weights are a better measure of case-level costs than the compressed weights are.

IMPLICATIONS FOR POLICY AND RECOMMENDATIONS

We suspect that the variation in costs across hospitals (especially between newer and older hospitals) is in part an artifact of TEFRA. Our empirical findings from the interim report included the fact that estimates of the effect of comorbidity are more precise if one controls for hospital costs using individual hospital identity results rather than standardized cost or standardized charges. The interim report also showed that the standard-method cost weights and the standard-method charge weights also have CMI compression. For this report, we therefore

updated the weight calculations--using the HSRV method chosen in the interim report (and in the NPRM), new data, and new definitions of CMG.

In the resulting weight regression, the effects of our comorbidity tiers are highly significant, very precise in large RICs, reasonably precise in all RICs, and almost always monotonic. The defects in monotonicity were corrected by averaging.

The HSRV weights, however, remain compressed. The improved FRG and comorbidity definitions have only a small effect on measured compression. We saw in the interim report that increasing the number of nodes also had only a minimal effect on compression. Consequently, we do not think that the hospitals with a high CMI are consistently getting patients who are more costly within each CMG and who are identifiable from their FIM scores, comorbidities, and/or age.

Our evidence strongly suggests that part of the problem arises from the bundling of variable ancillary services into the routine per diem. Insofar as this is true, the hospital cost-per-case estimates average out these errors in the patient-level cost estimates. Therefore it makes sense to "decompress" the weights using the formula in Eq. (6.3). This small adjustment to the weights results in only a small change--but it eliminates the relationship between the CMI and the hospital payment-to-cost ratio.

Table 6.6 shows the HSRV relative weights for each combination of CMG and comorbidity tier. These are the weights we used in the analyses reported in the rest of this report. The table also shows the average length of stay for cases paid as typical that we used to calculate the weights for transfer cases. The average LOS is based on 1999 data, but models were used to estimate the average when there were fewer than 10 cases in the cell.

Table 6.7 shows the sample sizes used in the calculation of relative comorbidity-adjusted weights. The first column is a count of the number of 1999 discharges after bundling. As discussed above, short-stay transfers are counted as only a fraction of a case, with the fraction equal to the LOS divided by the average LOS in the FRGC. The last column of Table 6.7 shows discharges using this counting method.

Table 6.8 shows the sample sizes used to calculate the comorbidity effects for the weights. Only cases discharged to the community and

long-stay transfers were used in this calculation. Although cases in RICs 50 and 51 are not included in the comorbidity calculations, their sample sizes are shown for completeness. Also for completeness, Table 6.9 provides the number of discharges and equivalent cases in each combination of tier and FRG. The small sample sizes in some tier-FRG combinations show why we did not attempt to obtain independent weights for these combinations.

Table 6.6

Weights and Expected Length of Stay, 1999 Data

CMG	Weights				Expected LOS			
	Tier 1	Tier 2	Tier 3	No comorb.	Tier 1	Tier 2	Tier 3	No comorb.
101	0.4778	0.4279	0.4078	0.3859	9.5	8.6	6.4	8.0
102	0.6506	0.5827	0.5553	0.5255	11.3	11.8	10.2	10.0
103	0.8296	0.7430	0.7080	0.6700	13.5	12.2	11.6	12.0
104	0.9007	0.8067	0.7687	0.7275	17.3	12.6	12.1	12.7
105	1.1339	1.0155	0.9677	0.9158	16.1	16.6	15.0	15.4
106	1.3951	1.2494	1.1905	1.1267	18.3	18.0	18.1	18.3
107	1.6159	1.4472	1.3790	1.3050	17.3	20.4	20.5	20.8
108	1.7477	1.5653	1.4915	1.4115	25.2	26.6	22.4	22.6
109	1.8901	1.6928	1.6130	1.5265	23.8	24.0	22.2	23.8
110	2.0275	1.8159	1.7303	1.6375	28.6	25.0	26.7	25.9
111	2.0889	1.8709	1.7827	1.6871	29.3	26.1	24.4	26.7
112	2.4782	2.2195	2.1149	2.0015	39.8	32.9	29.8	30.6
113	2.2375	2.0040	1.9095	1.8071	30.4	26.5	26.6	27.7
114	2.7302	2.4452	2.3300	2.2050	37.3	33.8	32.4	32.6
201	0.7689	0.7276	0.6724	0.6170	12.8	14.1	13.6	11.1
202	1.1181	1.0581	0.9778	0.8973	17.6	16.2	17.1	15.6
203	1.3077	1.2375	1.1436	1.0495	18.9	19.7	18.8	17.8
204	1.6534	1.5646	1.4459	1.3269	24.4	23.0	22.2	21.6
205	2.5100	2.3752	2.1949	2.0143	43.7	35.6	34.5	30.8
301	0.9655	0.8239	0.7895	0.7195	14.4	13.6	12.2	12.6
302	1.3678	1.1672	1.1184	1.0194	19.1	17.1	17.4	16.4
303	1.8752	1.6002	1.5334	1.3976	23.4	22.7	22.2	21.6
304	2.7911	2.3817	2.2824	2.0801	43.8	32.4	33.9	31.0
401	0.9282	0.8716	0.8222	0.6908	15.2	14.5	16.1	14.4
402	1.4211	1.3344	1.2588	1.0576	20.8	17.5	21.8	19.2
403	2.3485	2.2052	2.0802	1.7478	31.8	31.9	30.6	29.7
404	3.5227	3.3078	3.1203	2.6216	46.3	43.0	61.8	40.0
501	0.7590	0.6975	0.6230	0.5363	12.1	13.0	10.3	10.0
502	0.9458	0.8691	0.7763	0.6683	14.5	17.1	10.0	12.1
503	1.1613	1.0672	0.9533	0.8206	17.2	17.1	15.4	13.8
504	1.6759	1.5400	1.3757	1.1842	23.3	21.4	21.0	18.9
505	2.5314	2.3261	2.0778	1.7887	30.6	30.7	28.8	28.3
601	0.8794	0.6750	0.6609	0.5949	13.8	12.8	11.6	11.8
602	1.1979	0.9195	0.9003	0.8105	14.5	14.5	13.8	15.0
603	1.5368	1.1796	1.1550	1.0397	21.1	18.2	18.0	18.0
604	2.0045	1.5386	1.5065	1.3561	30.5	23.6	24.7	22.8
701	0.7015	0.7006	0.6710	0.5960	12.5	12.9	11.8	11.4
702	0.9264	0.9251	0.8861	0.7870	15.3	15.0	15.6	13.9
703	1.0977	1.0962	1.0500	0.9326	17.6	17.4	16.7	16.3
704	1.2488	1.2471	1.1945	1.0609	14.4	19.7	18.7	18.1
705	1.4760	1.4740	1.4119	1.2540	20.0	22.1	21.6	21.0
801	0.4909	0.4696	0.4518	0.3890	9.4	8.6	8.0	8.1
802	0.5667	0.5421	0.5216	0.4490	10.0	10.3	8.9	8.6
803	0.6956	0.6654	0.6402	0.5511	9.4	11.1	11.0	9.9
804	0.9284	0.8881	0.8545	0.7356	15.2	14.1	13.7	12.1
805	1.0027	0.9593	0.9229	0.7945	15.7	15.5	14.2	13.6
806	1.3681	1.3088	1.2592	1.0840	20.5	20.3	18.6	17.9
901	0.6988	0.6390	0.6025	0.5213	11.8	11.2	10.9	10.5

Table 6.6 (cont.)

CMG	Weights				Expected LOS			
	Tier 1	Tier 2	Tier 3	No comorb.	Tier 1	Tier 2	Tier 3	No comorb.
902	0.9496	0.8684	0.8187	0.7084	14.6	14.5	13.5	13.3
903	1.1987	1.0961	1.0334	0.8942	18.0	17.9	17.3	16.1
904	1.6272	1.4880	1.4029	1.2138	23.2	23.4	22.5	21.1
1001	0.7821	0.7821	0.7153	0.6523	13.1	13.3	11.6	13.0
1002	0.9998	0.9998	0.9144	0.8339	15.2	14.9	14.0	15.3
1003	1.2229	1.2229	1.1185	1.0200	17.7	17.2	16.7	17.6
1004	1.4264	1.4264	1.3046	1.1897	19.9	20.3	19.2	19.2
1005	1.7588	1.7588	1.6086	1.4670	21.1	24.6	23.0	23.4
1101	1.2621	0.7683	0.7149	0.6631	18.2	11.2	12.9	12.4
1102	1.9534	1.1892	1.1064	1.0263	24.5	18.1	16.9	17.5
1103	2.6543	1.6159	1.5034	1.3945	33.3	23.4	22.4	24.7
1201	0.7219	0.5429	0.5103	0.4596	12.6	10.0	10.7	9.4
1202	0.9284	0.6983	0.6563	0.5911	16.4	10.8	13.4	12.7
1203	1.0771	0.8101	0.7614	0.6858	17.6	14.6	13.7	13.3
1204	1.3950	1.0492	0.9861	0.8882	22.0	18.7	16.4	16.8
1205	1.7874	1.3443	1.2634	1.1380	26.7	21.0	20.9	20.4
1301	0.7719	0.6522	0.6434	0.5566	13.2	13.8	12.6	11.1
1302	0.9882	0.8349	0.8237	0.7126	15.7	13.9	14.0	13.7
1303	1.3132	1.1095	1.0945	0.9469	19.5	17.5	16.1	17.4
1304	1.8662	1.5768	1.5555	1.3457	24.9	24.8	29.3	22.3
1401	0.7190	0.6433	0.5722	0.5156	15.3	12.0	10.6	11.0
1402	0.9902	0.8858	0.7880	0.7101	12.5	14.5	13.4	13.4
1403	1.2975	1.1608	1.0325	0.9305	20.9	18.8	15.8	15.9
1404	1.8013	1.6115	1.4335	1.2918	30.3	24.0	21.4	20.3
1501	0.8032	0.7633	0.6926	0.6615	14.8	12.5	13.4	12.9
1502	1.0268	0.9758	0.8855	0.8457	17.0	17.0	13.6	15.2
1503	1.3242	1.2584	1.1419	1.0906	20.6	19.6	17.8	18.2
1504	2.0598	1.9575	1.7763	1.6965	30.1	28.3	30.0	26.3
1601	0.8707	0.8327	0.7886	0.6603	15.4	13.8	13.4	13.1
1602	1.3320	1.2739	1.2066	1.0103	21.2	19.8	19.9	17.6
1701	0.9996	0.9022	0.8138	0.7205	15.7	13.8	11.1	13.4
1702	1.4755	1.3317	1.2011	1.0634	21.0	20.5	20.2	18.1
1703	2.1370	1.9288	1.7396	1.5402	32.6	28.2	26.9	24.4
1801	0.7445	0.7445	0.6862	0.6282	12.0	12.0	11.5	10.0
1802	1.0674	1.0674	0.9838	0.9007	16.4	16.4	15.7	16.1
1803	1.6350	1.6350	1.5069	1.3797	22.4	25.1	19.5	21.6
1804	2.9140	2.9140	2.6858	2.4589	41.4	28.5	39.8	39.5
1901	1.1585	1.0002	0.9781	0.8876	15.3	15.0	16.4	15.0
1902	2.1542	1.8598	1.8188	1.6505	27.4	26.5	27.0	24.1
1903	3.1339	2.7056	2.6459	2.4011	41.1	34.7	29.5	39.9
2001	0.8371	0.7195	0.6705	0.6029	12.4	12.7	11.3	11.7
2002	1.1056	0.9502	0.8855	0.7962	15.1	15.2	14.2	14.2
2003	1.4639	1.2581	1.1725	1.0543	19.9	18.4	17.9	18.0
2004	1.7472	1.5017	1.3994	1.2583	29.6	22.2	21.3	22.1
2005	2.0799	1.7876	1.6659	1.4979	32.9	25.2	24.3	23.7
2101	1.0357	0.9425	0.8387	0.8387	17.5	17.5	14.6	16.2
2102	2.2508	2.0482	1.8226	1.8226	30.6	25.7	26.3	28.7
5001	.	.	.	0.1651	.	.	.	2.6
5101	.	.	.	0.4279	.	.	.	7.9
5102	.	.	.	1.2390	.	.	.	22.5
5103	.	.	.	0.5436	.	.	.	8.8
5104	.	.	.	1.7100		.	.	27.9

Table 6.7

Sample Sizes for Calculation of Relative CMG Weights (1999 data)

CMG	Discharges	Equivalent Cases
101	268	266
102	1,727	1,707
103	623	610
104	4,577	4,501
105	7,763	7,575
106	7,088	6,817
107	4,046	3,815
108	1,536	1,407
109	4,948	4,588
110	1,444	1,170
111	1,785	1,577
112	2,921	2,284
113	5,734	5,207
114	9,035	7,340
201	362	352
202	381	364
203	797	753
204	629	557
205	918	740
301	1,230	1197
302	1,468	1,398
303	1,703	1,544
304	930	745
401	240	234
402	375	359
403	482	419
404	318	245
501	1,611	1,589
502	266	262
503	2,368	2,301
504	1,232	1,154
505	1,963	1,676
601	1,354	1,334
602	2,976	2,890
603	3,845	3,644
604	3,594	3,193
701	3,841	3,787
702	6,067	5,893
703	4,563	4,366

Table 6.7 (cont.)

CMG	Discharges	Equivalent Cases
704	4,152	3,922
705	9,591	8,394
801	5,359	5,323
802	5,025	4,987
803	19,283	19,112
804	10,350	10,159
805	3,601	3,484
806	4,081	3,755
901	1,897	1,875
902	3,357	3,283
903	3,785	3,618
904	3,033	2,676
1001	497	489
1002	1,572	1,540
1003	1,768	1,700
1004	1,748	1,651
1005	2,583	2,240
1101	268	262
1102	433	401
1103	198	167
1201	670	666
1202	546	534
1203	1,818	1,788
1204	1,742	1,666
1205	1,327	1,208
1301	650	640
1302	746	734
1303	923	871
1304	645	575
1401	2,317	2,275
1402	3,274	3,145
1403	2,886	2,685
1404	1,702	1,461
1501	996	977
1502	2,907	2,796
1503	1,872	1,730
1504	1,177	958
1601	2,304	2,252
1602	1,358	1,251
1701	679	663
1702	974	925
1703	597	509
1801	35	35
1802	59	58

- 217 -

Table 6.7 (cont.)

CMG	Discharges	Equivalent Cases
1803	161	149
1804	115	90
1901	121	118
1902	172	165
1903	108	86
2001	5,386	5,252
2002	9,254	8,905
2003	9,118	8,444
2004	1,702	1,415
2005	3,262	2,747
2101	57	51
2102	96	86
5001	4,953	4,953
5101	96	96
5102	57	57
5103	654	654
5104	356	356

- 218 -

Table 6.8

Sample Sizes for Calculation of Comorbidity Adjustments to Weights
(combined 1998 and 1999 data)

| RIC | Number of Cases Paid as Typical Cases | | | | Total |
	Tier 1	Tier 2	Tier 3	No Relevant Comorbidity	
1	501	2,101	8,567	73,795	84,964
2	95	599	778	3,117	4,589
3	119	1,078	1,949	5,049	8,195
4	53	183	234	1,718	2,188
5	72	858	1,169	10,187	12,286
6	115	1,891	2,489	14,419	18,914
7	114	1,779	3,191	39,829	44,913
8	112	2,901	4,025	81,565	88,603
9	52	815	1,650	17,053	19,570
10	65	1,216	4,225	8,011	13,517
11	3	200	585	596	1,384
12	21	394	813	8,624	9,852
13	15	252	427	4,066	4,760
14	186	1,584	2,508	12,162	16,440
15	41	847	721	9,139	10,748
16	15	291	511	5,270	6,087
17	51	177	320	3,094	3,642
18	18	58	59	417	552
19	21	88	81	475	665
20	697	5,489	9,299	28,626	44,111
21	6	32	40	163	241
50	0	0	0	8,978	8,978
51	0	0	0	2,232	2,232

Table 6.9

Number of Cases and Equivalent Cases, by CMG and Tier
(combined 1998 and 1999 data)

FRG	Tier 1	Tier 2	Tier 3	No Relevant Comorbidity	Tier 1	Tier 2	Tier 3	No Relevant Comorbidity
101	1	8	45	568	1	8	44	563
102	17	44	291	3,468	17	43	285	3,426
103	1	17	88	1,212	1	17	86	1,188
104	26	146	793	8,425	26	142	767	8,279
105	54	306	1,507	13,580	51	285	1,445	13,258
106	66	327	1,427	12,242	64	309	1,347	11,766
107	42	216	828	6,839	41	200	762	6,476
108	12	66	273	2,756	11	56	248	2,518
109	71	280	1168	8,307	65	251	1,058	7,765
110	11	82	311	2,249	10	68	237	1,857
111	20	85	347	3,026	18	73	300	2,661
112	48	217	652	4,731	35	170	462	3,688
113	75	384	1,485	9,553	68	337	1,285	8,697
114	259	803	2,617	14,447	203	630	1,978	11,800
201	1	35	111	523	1	33	107	511
202	6	54	137	493	6	52	130	474
203	17	158	228	1,153	17	148	211	1,102
204	16	176	230	779	15	153	202	703
205	99	452	341	892	77	359	265	736
301	12	183	492	1,685	12	175	478	1,642
302	19	294	655	1,826	18	281	618	1,737
303	57	501	858	1,790	50	448	778	1,629
304	69	489	460	799	58	383	363	635
401	1	18	44	388	1	17	42	379
402	3	51	82	612	3	49	74	589
403	19	101	127	694	17	87	105	629
404	60	89	73	401	47	65	54	325
501	5	140	235	2,842	5	132	232	2,811
502	1	33	48	455	1	30	46	447
503	17	298	425	3,779	17	273	395	3,695
504	11	204	286	1,802	10	181	263	1,689
505	72	502	507	2,610	56	407	408	2,250
601	12	181	333	2,210	12	176	322	2,182
602	23	428	724	4,471	22	414	694	4,365
603	41	733	998	5,265	38	687	941	5,018
604	80	1,050	1,026	4,520	65	910	868	4,087
701	13	197	448	7,260	13	192	439	7,150
702	17	372	771	10,782	17	360	732	10,497
703	22	311	639	7,677	20	293	599	7,377

Table 6.9 (cont.)

FRG	Tier 1	Tier 2	Tier 3	No Relevant Comorbidity	Tier 1	Tier 2	Tier 3	No Relevant Comorbidity
704	22	337	615	6,742	20	316	556	6,369
705	76	1,127	1,691	14,873	67	960	1,417	13,056
801	9	240	362	10,883	9	236	357	10,821
802	6	257	358	9,515	5	251	350	9,454
803	41	1,086	1,559	34,906	41	1,062	1,515	34,621
804	31	809	1,128	17,242	31	781	1,079	16,940
805	12	299	389	6,121	11	286	363	5,955
806	24	533	723	6,399	21	476	635	5,888
901	10	127	259	3,498	10	123	253	3,455
902	18	212	484	5,388	16	200	466	5,285
903	16	334	622	6,051	16	308	581	5,785
904	18	399	646	4,383	16	334	553	3,886
1001	1	55	294	757	1	55	289	744
1002	7	241	953	2,056	7	232	930	2,014
1003	19	282	1,078	2,042	19	262	1,033	1,966
1004	15	326	1,062	1,921	15	297	1,002	1,806
1005	34	646	1,656	2,642	30	562	1,437	2,302
1101	.	55	212	204	.	54	206	200
1102	3	124	334	353	3	110	316	329
1103	1	92	159	160	1	77	134	131
1201	3	42	70	1,262	3	41	69	1,252
1202	1	38	81	871	1	38	79	852
1203	5	123	267	2,789	5	122	258	2,746
1204	8	136	284	2,589	7	124	264	2,500
1205	10	131	238	1,896	8	114	216	1,741
1301	3	46	81	1,076	3	45	79	1,060
1302	1	66	114	1,182	1	65	112	1,160
1303	7	96	174	1,384	6	89	161	1,318
1304	11	106	138	908	9	92	120	814
1401	20	301	521	3,747	20	289	507	3,667
1402	46	521	876	4,514	45	498	830	4,340
1403	59	607	919	3,626	57	553	843	3,407
1404	88	508	706	1,855	81	435	583	1,588
1501	7	79	75	1,848	7	76	71	1,818
1502	18	311	327	4,569	15	295	311	4,393
1503	18	314	268	2,671	16	290	249	2,456
1504	14	361	224	1,522	11	300	176	1,228
1601	8	153	362	3,891	7	146	348	3,815
1602	11	192	245	1,923	10	177	219	1,779
1701	9	50	100	1,200	8	48	96	1,180
1702	24	88	178	1,599	23	85	163	1,521
1703	27	76	156	848	24	67	120	731
1801	2	2	2	62	2	2	2	61
1802	2	12	18	89	2	12	18	86
1803	7	33	35	242	7	29	33	226

Table 6.9 (cont.)

FRG	Tier 1	Tier 2	Tier 3	No Relevant Comorbidity	Tier 1	Tier 2	Tier 3	No Relevant Comorbidity
1804	17	33	26	139	13	26	19	105
1901	5	18	31	193	5	17	30	189
1902	6	40	39	226	6	39	35	215
1903	12	58	36	137	11	48	28	110
2001	114	1,016	1,968	7,265	111	986	1,913	7,107
2002	197	1,874	3,495	10,897	185	1,775	3,338	10,509
2003	277	2,331	3,686	10,130	253	2,124	3,369	9,441
2004	62	515	583	1,872	48	418	477	1,580
2005	248	1,180	1,584	3,106	202	978	1,294	2,657
2101	0	7	20	81	0	7	17	79
2102	7	33	35	117	7	29	32	103
5001	0	0	0	8,978	0	0	0	8,978
5101	0	0	0	191	0	0	0	191
5102	0	0	0	100	0	0	0	100
5103	0	0	0	1,246	0	0	0	1,246
5104	0	0	0	695	0	0	0	695
All	3,113	29,109	54,656	386,783	2,761	26,348	49,619	366,647

7. FACILITY-LEVEL ADJUSTMENTS

OVERVIEW OF METHODOLOGY AND FINDINGS

In this section, we report on our analyses related to potential facility-level adjustments to the standard payment amounts. *Facility-level adjustments* are for systematic cost differences that are beyond the control of facility management and are appropriate to recognize in the payment system. In addition to the adjustment required by law for geographic differences in wage levels, our interim report recommended that adjustments be made for serving low-income patients and for location in a rural area (i.e., outside a Metropolitan Statistical Area, or MSA). This report updates the analyses in the interim report using more-recent data and improved measures for factors that may account for systematic cost differences.

Our methodology for evaluating the effects of various factors on a facility's costs per case involved several steps. First, we updated and refined the explanatory variables that we used in our preliminary analyses for the interim report. Next, we performed multivariate regression analyses to measure the effects of the factors on facility costs and to establish potential payment adjustments. Third, we confirmed the adjustments indicated by the regression analyses by means of payment simulations. Finally, we undertook supplemental analyses to understand some of the regression and simulation results.

Our analysis file consists of 714 facilities for which we have case mix and cost data (compared to 624 in the interim report). For most facilities, we used the HCRIS 14 and 15 cost reports to establish the facility-level variables for our regression analyses and payment simulations.

In the first subsection below, we describe the variables that we used in our regression analyses--generally, the same variables that we used in the interim report. The potential payment variables included updated measures for the area wage adjustment, location in a large urban or rural area, the indirect costs of graduate medical education (IME), and a variable for serving low-income patients. We also established

updated measures for other factors that may explain costs such as the date of certification, type of facility, and type of ownership. We refined some measures based on comments from the TEP and our review of the methodologies that we used in the interim report.

In the second subsection, we describe our methodology for evaluating the effect of the explanatory variables on facility costs. We sought to improve the fit and predictive power of the facility-level model developed in the interim report. Our primary tool was multivariate regression. The dependent variable is the average cost per case at a particular facility. The general specification is

$$C = f\ (CMI,\ WI,\ X)$$

where:

C = average cost per case at the facility

CMI = the case mix index, a measure of the relative resources required to treat cases at the facility

WI = the hospital wage index for the facility, a measure of relative differences in input prices

X = a vector of additional explanatory variables that may affect a hospital's costs per case, such as its teaching activities, proportion of low-income patients, etc.

We performed our analysis in several steps. First, we used fully specified regressions to understand the various factors affecting cost and to identify which potential payment variables were significant. Next, we performed a set of regressions using only the potential payment variables. We determined which variables were significant and explored the effects of using different specifications and measures for those variables on their coefficients.

We used the coefficients from the payment regressions as potential payment adjustment factors. We then simulated payments using the 1999 cases in our analysis file and determined the PTC ratios for different classes of hospitals for specific combinations of payment policies. Finally, we undertook supplemental analyses to investigate some of the regression results that differed from the findings in our interim report. For example, we examined the effects of individual hospital

outliers on our regression results using a statistical technique called "added variable plots" as well as additional regressions.

In the third subsection, we present our results. Key findings include the following:

- In the fully specified regressions using a comprehensive set of explanatory variables, the following factors were significant in explaining variation in cost per case: case mix, wage index, the proportion of low-income patients, type of facility (freestanding hospital or unit), certification date, size, and geographic location. In the interim report, teaching and type of ownership were also significant in the fully specified regression.

- Teaching was not significant in either the fully specified or payment regressions using different measures of teaching intensity. However, problems in measuring interns and residents may have affected our results.

- The wage index was compressed when we defined it as an independent variable. Compression was not evident in the interim report and may be attributable to a small number of outlier hospitals.

- When cost per case was standardized for case mix and area wage differences, rural hospitals were 19 percent more costly than other hospitals. Large urban hospitals were almost 4 percent less costly than other hospitals.

- There was about a 5 percent increase in costs for each 10 percentage point increase in a facility's disproportionate share percentage of low-income patients.

In the final subsection, we discuss the implications of our findings for policy and our recommendations.

ESTABLISHING THE VARIABLES USED IN THE ANALYSES

Table 7.1 summarizes the differences in the data sources and measures that we used in the analyses for this report compared to those we used in our preliminary analyses for the interim report. We discuss the changes in more detail below.

- 226 -

Table 7.1

Comparison of Data Used in Interim and Final Reports

Data Element	Interim Report	Final Report
Claims data	CY 1996 and 1997	CY 1998 and 1999 1998 and 1999: 647 facilities 1998 only: 20; 1999 only: 47
Cost reports	HCRIS 13-14--latest files available as of 12/99; HCRIS-12 if necessary	HCRIS 14-15--latest files available as of 4/2000; HCRIS 13 if necessary HCRIS 14 and 15: 618 facilities HCRIS 13 and 14: 86 HCRIS 13 and 15: 2 HCRIS 13 or 14 or 15 only: 8
Average cost per case	Case-weighted average for 1996 and 1997 cases	Case-weighted average for 1998 and 1999 cases
Average case mix index	Case-weighted average for 1996 and 1997 cases	Case-weighted average for 1998 and 1999 cases
Wage index	FY 2000 hospital wage index (from HCRIS 13 data): non-reclassified; no GME wages	FY 2001 hospital wage index (from HCRIS-14 data): non-reclassified; no GME wages
SSI percentage	Derived from FY 1998 MEDPAR claims for the total hospital (acute and rehab unit)	Derived from FY 1999 MEDPAR claims for rehabilitation facility (hospital or unit)
Medicaid percentage	Latest available cost report; imputed missing low-income values based on statewide average	Latest available cost report; used SSI percentage and state to predict missing values
Resident salary allocation for FTE count	Latest cost report available as of 10/99 (mix of HCRIS 13 and 14)	HCRIS 14 and HCRIS 15 cost reports available as of 12/00
Resident to average daily census ratio	Resident count and average daily census from latest available cost report(s)	Ratio from latest available cost reports; match the cost report used to allocate salaries
Certification date	OSCAR--10/99 version	Date of certification provided on HCRIS 12 cost report; if missing, earlier of first claims year or OSCAR date

Note: OSCAR = Online Survey and Certification Reporting system.

Cost per Case

We defined a *case* in the manner consistent with the policies that are being implemented in the IRF PPS (discussed earlier in this report). In particular, short-stay transfers to another hospital, SNF, or nursing facility were counted as a partial discharge, and interrupted stays were bundled together into a single discharge. A short-stay transfer case's equivalence to a full case was determined by the ratio of the length of stay for the transfer plus an additional half-day to the average length of stay for all typical cases in the same CMG. Two or more discharges counted as a single case when a patient was discharged from the rehabilitation facility and returned to the same facility on the day of discharge or either of the following two calendar days.

We determined the facility's average cost per case by summing the costs for all cases in the analysis sample and dividing by the number of equivalent full cases. When we had cost-per-case data for both 1998 and 1999, we pooled the number of cases and total costs. The pooled data should provide more stability in the payment adjustments than data for a single year. We did not adjust for cost differences between the years for which we pooled the data. The case-weighted average increase in cost per case between 1998 and 1999 for a matched set of facilities was less than .2 percent. When data for only one year were available, we used the costs and number of equivalent cases for that year. This allowed us to include the maximum number of facilities in our analysis having a case weight consistent with each facility's share of Medicare cases.

We used the cost per case calculated from the analysis file rather than the cost per discharge from the cost report. This (1) provided a match between the cases for which we have case mix data and the cost of those cases and (2) accounted for transfer cases and interrupted stays. By treating short-stay transfers as partial discharges and bundling interrupted stays, the dependent variable is made consistent with the final payment policies under the IRF PPS.

Case Mix Index

The *case mix index* (CMI) is the average of the CMG relative weights derived by the HSRV method for each facility. We normalized the case mix index to 1.0 for complete cases. We gave short-stay transfers a partial

weight based on the ratio of the length of stay for the transfer plus one-half day to the average length of stay for all typical cases. If we had cases for both 1998 and 1999, we computed a case-weighted average CMI for the two years. We computed two CMI values for each facility: one using the compressed relative weights and one using the decompressed weights (see Section 6).

Wage Index Value

The wage index adjustment is intended to account for systematic differences in wage levels across labor market areas. In the interim report, we used a CMS-furnished wage index based on the FY 1996 audited hospital wage data. For this report, we used a CMS-furnished wage index based on the FY 1997 audited hospital wage data, that is, wage data for cost reporting periods beginning on or after October 1, 1996 and before October 1, 1997. The labor market areas are consistent with those in other Medicare prospective payment systems (i.e., MSAs and non-MSA areas of states) but are determined without regard to hospital geographic reclassification under section 1866(d)(8) or (d)(10) of the Medicare law. As in the interim report, the wage index data excluded 100 percent of wages for services provided by teaching physicians, interns and residents, and non-physician anesthetists under Part B.

For the interim report, we used fully specified regressions to examine alternative specifications for the wage index variable. The results indicated that a wage index variable using a predetermined labor-related share performed as well as specifications using the coefficient of wage index variable to establish the wage index adjustment factor. Since CMS expressed a preference for using the predetermined labor-related share in the IRF PPS, we used it to develop the payment adjustments in the interim report. We used the same specification for this report. We defined the wage index variable as $(.705 * WI + .295)$. This is consistent with the way the wage index would be applied in an IRF PPS using a 70.5 percent labor-related share. The

70.5 percent is the CMS Office of the Actuary estimate for the labor-related share of TEFRA hospital costs in FY 1998.[1]

Geographic Location

As in the interim report, we established variables to identify whether a hospital is located in a large urban area, another urban area, or a rural area.

The legislation establishing the IRF PPS specifically authorizes a cost-of-living adjustment for rehabilitation facilities in Alaska and Hawaii. There are three rehabilitation facilities in Alaska and one in Hawaii. However, our analysis file contains only one of the Alaskan facilities and the Hawaiian hospital. We examined the impact of the IRF PPS on these hospitals in our payment simulations.

Indirect Teaching Costs

Evaluation of the effect of teaching on rehabilitation facility costs is complicated by differences in reporting resident counts for freestanding rehabilitation hospitals and units. No full-time equivalent (FTE) resident count specific to inpatient rehabilitation services is reported. For freestanding hospitals, the resident count for the hospital includes resident time spent on outpatient as well as inpatient services. For rehabilitation units, a resident count for the routine area of the rehabilitation unit is reported. The count should not include the time residents spend in ancillary departments furnishing services to rehabilitation inpatients. This time should be assigned to the ancillary department, where it is combined with resident time attributable to ancillary services furnished to patients in the acute care portion of the hospital and to outpatients.

In developing a measure of teaching intensity for the interim report, we assumed that the FTE resident count reported for the

---

[1] We used the 70.5 percent because it matches our cost report and cost-per-case data. In our FY 2002 payment simulations, we used the CMS estimate of the FY 2002 labor-related share (72.40) to be consistent with the payment policies that are used under the IRF PPS. In the interim report, we determined the wage data for Albany, Georgia, was aberrant. However, the wage index value for that city appears reasonable in the updated wage data.

rehabilitation unit represents the time residents are assigned to the rehabilitation unit. We assumed it does not include time residents in physical medicine and rehabilitation programs spend in ancillary and outpatient departments. A comparable FTE count for freestanding rehabilitation hospitals would be resident time spent in the inpatient routine areas. This FTE count, however, is not reported. To develop a consistent count, we estimated the number of residents assigned to the routine areas of freestanding hospitals based on the ratio of resident salaries apportioned to those areas to total resident salaries for the facility.

The data that we needed to develop the estimates (FTE count for the rehabilitation unit; percentage of resident salaries assigned to routine areas of freestanding hospitals) are reported to CMS on the electronic cost reporting system but are not available on the CMS public use file. For the interim report, CMS provided this information for the latest available cost reports as of October 1999 (a combination of HCRIS 13 and HCRIS 14 cost reports. Data on the distribution of resident salaries are on the cost report for informational purposes only because Medicare's direct GME payments are based on the program's share of a prospectively determined per-resident amount. As a result, the data are not audited. Similarly, the number of residents assigned to the rehabilitation unit (reported on Worksheet S-3) does not affect payment for rehabilitation services. These data elements are not likely to be audited and may not be as reliable as data elements that directly affect payment. (If anything, acute care hospitals would have an incentive to minimize the resident count for the rehabilitation unit since it is not included in the indirect medical education adjustment to the DRG payment for acute care inpatient services.)

The TEP noted that the ratio of residents to average daily census (ADC) is not consistent with the Residency Review Committee (RRC) accreditation requirements for programs in physical medicine and rehabilitation and suggested that there may be reporting errors. Two RRC requirements are of particular interest (AMA, 1999). The first states that a minimum of eight inpatients should be available for each resident assigned full-time to an inpatient rehabilitation service. Using this as a yardstick, we would expect the maximum ratio of residents to average

daily census would be .125 unless the residents are assigned to care for only a subset of the facility inpatients. In our analysis file for the interim report (624 facilities), there were 15 facilities with a resident-to-ADC of .20 or higher. Fourteen of the 15 facilities were units of acute care hospitals, where our FTE count was based on the number reported on the cost report for the rehabilitation unit--a finding inconsistent with the acute care PPS incentive to minimize the resident count for such units. The second standard requires that during three years of the physical medicine and rehabilitation residency program, a resident spend at least one-third of the time in the management of inpatients with a need for physical medicine and rehabilitation and at least one-third of the time in outpatient care.

We took several steps to investigate this issue further using data from the interim report analysis file. First, we tested whether the number of residents reported for the rehabilitation units was consistent with the proportion of resident salaries allocated to those units. We found that there were significant differences between the FTE resident counts for some units. Table 7.2 gives the results for the 12 units reporting five or more residents. (All but one had a resident-to-ADC ratio greater than .125.) The resident count in column (a) is based on the total number of residents reported for the hospital complex. Column (b) shows the number of residents reported for the rehabilitation unit.[2] Column (c) shows an estimated resident count determined by multiplying the FTE count for the hospital complex by the ratio of resident salaries assigned to the rehabilitation unit to total resident salaries for the facility.[3] There is less than a 1.0 difference in the resident count for three of the 12 facilities. Compared to the reported FTE count, the estimated count is substantially higher for three units (Units C, D and F) and substantially lower for the remaining seven units. It is not readily apparent which measure is the more reliable estimate of resident time on the rehabilitation unit.

---

[2] Worksheet S-3, Part I, Line 14, Column 9.

[3] Worksheet S-3, Part I, Line 25, Column 9 x Worksheet B, Part I, Line 31 for rehabilitation sub-provider, Column 22 / Worksheet B, Part I, Line 22, Column 22.

Table 7.2

Comparison of FTE Resident Counts Using Different Measures

Rehab. Unit	(a) Facility FTE Residents	(b) Reported Rehab. Unit FTE Residents	(c) Estimated Rehab. Unit FTE Residents	(d) Estimated Total Rehab. FTE Residents	(e) Resident-to-ADC Ratio Based on Reported FTEs
A	271.08	6.55	6.41	7.78	.1758
B	252.85	5.80	6.69	6.89	.1665
C	32.81	9.07	12.54	12.54	.1309
D	588.23	6.00	22.99	23.66	.2536
E	49.35	5.17	5.36	5.79	.2381
F	168.67	7.75	23.38	23.38	.4169
G	90.33	8.90	4.42	4.79	.3795
H	282.52	10.87	6.74	7.32	.1088
I	521.82	11.02	6.82	7.42	.5612
J	470.32	13.27	7.97	8.76	.4256
K	303.99	9.83	3.95	4.43	.6229
L	221.89	10.59	3.64	7.08	.3952

Second, we developed an estimated FTE resident count that would take into account all resident time for inpatient rehabilitation (routine and ancillary services). We based the estimate on the proportion of resident salaries that is attributable to inpatient rehabilitation patients.[4] There is little or no impact on the percentage of resident salaries associated with rehabilitation inpatients of units of acute care hospitals. For the 141 units for which the salary data were available, the average percentage of resident salaries associated with all rehabilitation services is 6.2 percent, compared to 5.8 percent for routine services only. However, because the units are relatively small, a minor change in the resident count may have a noticeable effect on the measure of teaching intensity. We show the results for units with more than five residents in column (d) of Table 7.2.

---

[4] For freestanding hospitals, the ratio of inpatient charges to total charges can be used to estimate resident time associated with inpatient rehabilitation services. The estimate is more difficult for units. It requires computing an FTE resident count per Medicare inpatient day (based on Medicare rehabilitation inpatient charges to total charges in the departments where residents are assigned) and assuming that usage of ancillary services by Medicare patients is representative of usage by non-Medicare patients.

We found that 16 of the 25 freestanding hospitals for which the
resident salaries are available assigned 100 percent of resident
salaries to routine services. For the remaining nine hospitals, the
average percentage of resident salaries associated with all inpatient
rehabilitation services is 79 percent compared to 65 percent for routine
services only. One might expect that hospitals with only a few rotating
residents would assign their residents solely to the routine areas.
Given the RRC requirements, we would expect that hospitals with a larger
number of residents would have at least some resident time assigned to
outpatient services. However, as shown in Table 7.3, the salary
allocation pattern does not seem to be related to the number of
residents assigned to the facility.

Table 7.3

**Comparison of Proportion of Resident Salaries in Freestanding Hospitals
Allocated to Routine and Total Inpatient Rehabilitation Services**

Hos-pital	Facility FTE Residents	Rehab. Routine Salary (%)	Total Inpatient Rehab. Salary (%)	Hos-pital	Facility FTE Residents	Rehab. Routine Salary (%)	Total Inpatient Rehab. Salary (%)
A	1.25	100.0	100.0	L	12.52	76.7	83.6
B	1.00	100.0	100.0	M	11.00	100.0	100.0
C	13.11	54.0	57.3	N	1.75	100.0	100.0
D	11.17	100.0	100.0	O	3.55	87.8	87.8
E	26.62	63.1	78.7	P	9.76	50.0	50.4
F	14.31	100.0	100.0	Q	1.83	100.0	100.0
G	3.90	100.0	100.0	R	3.12	60.0	93.6
H	3.00	52.4	82.2	S	3.69	100.0	100.0
I	3.50	59.0	83.3	T	9.00	100.0	100.0
J	0.40	100.0	100.0	U	5.40	82.0	93.8
K	0.40	100.0	100.0	V	2.00	100.0	100.0
L	12.52	76.7	83.6	W	2.83	100.0	100.0
M	11.00	100.0	100.0	X	2.00	100.0	100.0
N	1.75	100.0	100.0	Y	0.97	100.0	100.0

We also examined how stable the resident counts and resident-to-ADC
ratios are from year to year. Our final analysis file of 714 facilities
contains 60 units and 17 freestanding hospitals that had a resident-to-

ADC ratio greater than 0 in both their HCRIS 14 and HCRIS 15 cost reporting periods.[5] In Table 7.4, we report the results of our analysis

Table 7.4

Changes in Resident Counts and Average Daily Census Between HCRIS 14 and HCRIS 15 Cost Reporting Periods for Matched Set of Facilities

	Freestanding Hospitals	Units of Acute Care Hospitals
Number with resident-ADC ratio > 0 in both HCRIS 14 and 15	17	60
Average FTE residents (facility-weighted)		
HCRIS 14	4.86	2.41
HCRIS 15	4.64	2.44
Average percentage change	21.7	39.5
N facilities with change > 1.0 FTE	4	18
N facilities with change > 10%	10	43
Average daily census (facility-weighted)		
HCRIS 14	78.3	28.3
HCRIS 15	80.3	29.3
Average percentage change in ADC	6.2	9.7
N facilities with change > 10%	3	19
Resident-to-ADC ratio		
HCRIS 14: average (facility-weighted) range	.0551 .0068-.1765	.1116 .0007-.6755
HCRIS 15: average (facility-weighted) range	.0586 .0068-.2200	.1225 .0038-.7593
Average percentage change	18.5	48.7
N facilities with change > 10%	9	44
N with residents > 0 in HCRIS 14 and = 0 in HCRIS 15	0	16
N with residents = 0 in HCRIS 14 and > 0 in HCRIS 15	1	6

---

[5] Fourteen units had a positive resident-to-ADC ratio in HCRIS 14 but no residents in HCRIS 15. Conversely, six units and one freestanding hospital had residents in HCRIS 15 but not HCRIS 14. In addition to these facilities, other facilities had incomplete data for either the HCRIS 14 or HCRIS 15 cost reporting periods. We have resident counts but not complete HCRIS 15 cost reports for an additional 14 rehabilitation units of freestanding hospitals.

[6] For this analysis, we used the resident counts reported for freestanding hospitals rather than the estimated FTE count for the routine area of freestanding hospitals because we were missing the salary allocation data for some hospitals for at least one of the cost reporting periods. Using the total hospital count allowed us to include these hospitals in the analysis.

of the resident counts for the matched set of 77 facilities having a
resident-to-ADC ratio greater than 0 in both periods.[6] The results show
that there was more than a 10 percent change in the resident-to-ADC
ratio for about half of the hospitals and two-thirds of the units. The
average percentage change for the units was nearly 50 percent compared
to 18 percent for the hospitals. Since units tend to have fewer
patients, their resident-to-ADC ratio is likely to be more sensitive to
changes in resident counts and patient census. On average, a 1.0
increase in the resident count for freestanding teaching hospitals
results in a .01 increment in the resident-to-ADC ratio for freestanding
hospitals (1:79) and a .03 increment in the ratio for units (1:29). Only
four hospitals and 18 units had changes of 1.0 FTE resident or more, and
most changes in the absolute values for the resident-to-bed ratios were
relatively small.

In summary, our analysis found the following:

- Salary allocations (and the resulting FTE resident counts that
  would result from using those allocations) are often inconsistent
  with residents' assigned time.
- The number of residents is quite small. However, because patient
  census and bed capacity are also relatively small, the measure of
  teaching intensity, particularly for units, is sensitive to
  relatively minor changes in resident counts.

For this report, we used the same approach as we used in the
interim report to calculate the number of FTE residents: For the
rehabilitation unit of acute care hospitals, we used the number of
residents reported for the unit;[7] for freestanding hospitals, we
estimated the number of residents assigned to the routine area based on
the ratio of resident salaries apportioned to those areas to total
resident salaries for the facility. We defined *teaching intensity* as the
ratio of FTE residents to ADC.

We decided not to evaluate in our regressions the FTE count
associated with all inpatient rehabilitation services. The differences
in resident counts between routine and total inpatient services for most

---

[7] CMS provided us with this information. It is not available as
part of the electronic cost reporting system.

facilities are minor based on current reporting practices. Given the reporting inconsistencies, we could not assess whether the more refined measure would sufficiently improve the explanatory power of the teaching intensity measure to justify the administrative burden that would be associated with determining the FTE count using this methodology on an ongoing basis.

In the interim report, we also tested a variable that indicated whether or not a facility had residents. For this report, we created a new measure of teaching commitment that does not rely on resident count. The measure is based on whether the facility is a major participating institution in a physical medicine and rehabilitation residency program.[8] We used the *Graduate Medical Education Directory, 1998-1999* (CD-ROM version) to determine which facilities are major participating institutions in physical medicine and rehabilitation residency programs, and we used this measure to distinguish between facilities that have occasional residents and those with regular ongoing resident rotations.

Low-Income Patients

In our interim report, we investigated an adjustment for serving low-income patients based on the facility's Medicare patients who are entitled to SSI and its non-Medicare patients who are entitled to Medicaid. We investigated two measures:

- We defined the *low-income patient (LIP) ratio* as the percentage of all inpatients who are either Medicare patients entitled to SSI or non-Medicare patients entitled to Medicaid.

- We defined the *disproportionate share of low-income patient (DSH) ratio* as the percentage of Medicare patients who are entitled to SSI plus the percentage of all inpatients who are eligible for Medicaid (but not Medicare). Under the PPS for acute care hospitals, this measure is used to make additional payments to hospitals serving a disproportionate share of low-income patients.

---

[8] As defined by the Accreditation Council for Graduate Medical Education, a major participating institution is one to which residents rotate for a required experience or which require explicit approval by the Residency Review Committee. The RRC for physical medicine and rehabilitation requires approval of any institution providing six months or more of training over the course of the residency.

In our interim report, we concluded that neither measure for low-income population appears to work better than the other. We also noted that we believed the low-income patient ratio is a better technical measure of low-income patient load than the DSH ratio is. However, to be consistent with the low-income patient measure for acute care hospitals, CMS proposed using the DSH ratio to determine the IRF PPS. We continued to evaluate both measures for this report.

Several measurement issues affecting both the LIP and DSH ratios emerged from the interim report that we discuss in this subsection. These issues are the availability of SSI information specific to rehabilitation stays, the methodology for imputing missing values, the effect that variation in Medicaid coverage policies might have on Medicare costs, and the accuracy of the Medicaid days reported on the Medicare cost reports.

SSI Percentage

We based the SSI measure in the interim report on the SSI percentage for all Medicare inpatient hospital stays. For units of acute care hospitals, the SSI percentage was for the acute care hospital rather than the rehabilitation unit. For the final report, CMS provided data on the SSI percentage specific to rehabilitation inpatient discharges occurring during 1999. This gave us an accurate measure of the proportion of Medicare days in the rehabilitation unit attributable to beneficiaries who are entitled to SSI.

Imputing Missing Values

In our interim report, we used a two-step process to impute missing values for our low-income patient measures:

- For rehabilitation units where we were missing only the Medicaid days, we estimated the Medicaid rehabilitation days by applying the ratio of Medicaid acute care days to total acute care inpatient days to the total inpatient rehabilitation days.
- If we were missing the SSI days or were also missing Medicaid days for the hospital, we imputed the low-income values by assigning the average LIP or DSH ratio for large urban facilities in the state or for other facilities in the state as appropriate.

We refined our methodology for imputing low-income values in the final report. Our regression analyses for the interim report indicated

that the facilities with missing values were significantly different from other facilities. Therefore, we investigated using the results from different prediction models to refine our imputations.

We predicted the proportion of non-Medicare days in the rehabilitation facility that are attributable to Medicaid patients as a function of the facility's percentage of Medicare patients who are entitled to SSI and the state in which the facility is located:

- We based the prediction model on the relationship in the state between SSI and the percentage of non-Medicare days attributable to Medicaid if we had SSI and Medicaid data available for more than 15 facilities in the state.

- If we had data for 15 or fewer facilities in the state, we based the prediction model on the national relationship between SSI and the percentage of non-Medicare patients entitled to Medicaid.

We decided on this prediction model after evaluating several alternatives that could not be applied to all the facilities in our analysis file and would have required a two-tiered imputation methodology.[9] The prediction model based on the facility's SSI percentage has the advantage of being derived from public information that is available for all rehabilitation facilities. The interaction of the SSI variable with the state variable takes into account both the facility's commitment to serving low-income patients and the expansiveness of the state's Medicaid eligibility and coverage of inpatient hospital rehabilitation services. We used all facilities for which we have both SSI data and Medicaid cost report data (including those that are not in our analysis file) in developing the prediction model. We then applied the results to 178 facilities in our analysis file.[10]

---

[9] The alternatives included using a measure of Medicaid utilization for the rehabilitation facility that we developed from UDSmr data. We did not have information on all patients that would allow us to develop a comparable measure for the HealthSouth facilities. We also explored using a model that took into account the Medicaid utilization for the acute care portion of the hospital (which would have required a different model for freestanding hospitals). The SSI prediction model performed nearly as well as the first alternative and better than the second.

[10] This includes 41 facilities with missing data and 137 facilities that reported 0 Medicaid utilization on their Medicare cost report. As

Accuracy of Reported Medicaid Utilization

Another issue (which was not raised by the TEP) is whether Medicaid days are accurately reported on the Medicare cost report. There are two areas of particular concern, both of which involve potential underreporting of Medicaid utilization. We assume that reporting accuracy will improve if the final payment system includes a facility adjustment for serving low-income patients. In the interim, however, the adjustment factors and budget neutrality could be affected if the Medicaid utilization rate used to determine the low-income patient adjustment is understated.

Our first concern was whether the Medicaid utilization rates of 0 reported in the cost report data for 198 facilities are actual values or missing values. We were able to investigate this issue using the data we received from UDSmr. It has information on the expected primary and secondary payer for all patients in each facility. For the facilities in our cost report file that are in the UDSmr database, we matched the days on the UDSmr bills to the cost reporting period in which they occurred after excluding days attributable to the acute portions of interrupted stays. We determined the proportion of inpatient days for which Medicaid is reported as the expected primary payment source. We computed Medicaid utilization rates and compared them with the rates reported on the cost report.

We found that 78 percent of the facilities with 0 Medicaid utilization on their Medicare cost report have Medicaid days reported in the UDSmr data for the same time period. When we used the two different values for Medicaid utilization in regression equations predicting cost per case, we found that the equation using the UDSmr data had the higher explanatory power. Based on this analysis, we decided to treat 0 Medicaid utilization rates as missing values when the utilization rate in the FIM data was greater than 0 or missing. We used the same methodology to predict the Medicaid utilization rate for these facilities that we used for other facilities with missing Medicaid

---

explained in the next subsection on accuracy of the Medicaid utilization data, we assigned Medicaid utilization as missing when its reported value was 0 on the cost report and was greater than 0 or missing in the FIM data.

utilization values. We did not assign a missing value when Medicaid
utilization was 0 in both the cost report and FIM (UDSmr) data.

A second area of concern was whether Health Maintenance
Organization (HMO) days were included in the Medicaid days reported for
rehabilitation units on the Medicare cost report. The PPS 15 cost
reports do not explicitly request Medicaid HMO days for subproviders. If
these days were not reported, the percentage of low-income patients
served by rehabilitation facilities would be understated, particularly
in areas with high Medicaid managed care penetration.

Variation in Medicaid Coverage Policies

The TEP expressed concern that there are significant differences in
state practices regarding coverage for Medicaid rehabilitation stays.
These differences are not a concern for the IRF PPS if the proportion of
Medicare patients who are entitled to SSI and the non-Medicare patients
who are entitled to Medicaid in the rehabilitation facility reflects the
care provided to low-income patients and its impact on the costs.
However, expected costs may be understated if charity care is
proportionately higher in those states with less-expansive Medicaid
coverage. Thus, our concern is less with lower Medicaid utilization
rates per se than with the possibility that the lower rates may
translate into higher charity care days that are not measured in the LIP
variable but that affect costs per case.

We used the UDSmr (Medicare and non-Medicare) data on primary and
secondary payer to investigate this issue further. We were limited to
these facilities because the HealthSouth data provide information on
Medicare patients only. Using the CMS-furnished data, we determined the
percentage of Medicare days attributable to patients who are also
entitled to SSI. From the UDSmr data, we constructed a measure of the
percentage of non-Medicare days attributable to Medicaid patients. We
also constructed two new measures related to LIP: the percentage of
Medicare days attributable to patients who are also entitled to Medicaid
("dual eligibles") and the percentage of non-Medicare patient days that
are self-pay or charity care.[11] In addition to the individual measures

---

[11] We used both charity care and self-pay patients in our low-
income measure because many hospitals do not distinguish between charity

for different types of low-income patients, we constructed an overall measure by adding the estimated Medicare dual eligible days to the Medicaid and charity care/self-pay days. With respect to the self-pay and charity care patient loads, we note that our analysis was limited to investigating the effects that serving low-income patients may have on Medicare cost per case. We did not intend to examine issues related to Medicare sharing in uncompensated care costs.

The SSI eligibility standards are based on federal poverty guidelines and do not vary based on differences in cost of living across geographic areas. As a result, the SSI measure tends to understate the percentage of poor Medicare beneficiaries in states with a relatively high cost of living. One advantage of the SSI measure is that it is not affected by the generosity of state Medicaid program policies pertaining to eligibility and coverage. The percentage of Medicare patients who are dually entitled to both Medicare and Medicaid also should not be affected by state coverage policies for inpatient rehabilitation stays, because Medicare is the primary payer for these patients. However, the dual eligible measure would reflect state differences in Medicaid eligibility rules affecting the Medicare population. For this reason, using the dual eligible ratio as a low-income measure may produce different results than the SSI ratio for Medicare patients.

Transformation of the Low-Income Patient Variable

The values for teaching and LIP ratios can be zero.[12] Since the log of zero cannot be taken, customary practice has been to add 1.0 to the ratios. This specification for the logged form of the variable is used in the acute care PPS for operating costs and was used to examine

---

care (for which the patient is not expected to pay from the time of admission) and bad debt (for which a patient's liability for at least some portion of the bill is waived upon determination of inability to pay). A high proportion of self-pay (or uninsured) bills become bad debt. A recent study of Massachusetts hospitals reported that most patients who incurred bad debt had incomes below the poverty line (Weissman, Dryfoos, and London, 1999).

[12] This occurs with the teaching variable for all non-teaching facilities. It rarely occurs with the LIP variable because at least one patient is likely to be entitled to SSI and/or Medicaid. Since teaching was not significant, our interim report recommendation relied on an analysis of the DSH adjustment.

teaching and DSH effects in the proposed rule for the hospital outpatient PPS. Rogowski and Newhouse (1992) found that, if the true specification is log-log, adding 1.0 to the measure of teaching intensity biases the teaching coefficient substantially. To reduce the distortion to less than 1 percent, they added 0.0001 to the teaching ratio instead of 1.0. Others have used the non-logged form (O'Dougherty, et al., 1992). More recently, Dalton and Norton (2000) concluded that HCFA's original specification of the teaching variable (1 + resident-to-bed ratio) is supported by the data.

In our interim report, we explored three different transformations of the low-income variable: a non-logged form of the variable and two logged forms (i.e., log (1.0 + DSH) and log (.0001 + DSH)). Our recommendation was to use the logged form that adds .0001 to the patient ratio. We found that while the differences were slight, this form provided a slightly better fit than either the non-logged form or the log (1.0 + patient ratio) form. However, the log (.0001 + patient ratio) form results in large differences in adjustment factor among hospitals with relatively low DSH ratios and modest differences for hospitals with relatively higher DSH ratios. For this report, we reexamined both log and non-logged transformations and the implications of using the results as the basis for a payment adjustment.

Other Factors Affecting Cost

In addition to examining factors that may be appropriate to incorporate into the IRF PPS, we explored the effect of other factors on rehabilitation facility costs. We used the same variables as we used in the interim report.

Type of Facility

There are 169 freestanding hospitals and 545 rehabilitation units of acute care hospitals in our analysis file. We used a dummy variable to evaluate whether there are systematic differences in costs between freestanding rehabilitation hospitals and rehabilitation units of acute care hospitals. In the interim report, the fully specified regressions found that freestanding hospitals were significantly more expensive than units of acute care hospitals.

Size

We used average daily census as the size variable and controlled for type of facility. Table 7.5 shows the dummy variables that we used for the size categories. When we used these variables in the fully specified regressions in the interim report, we found that small facilities were significantly more expensive than others.

Table 7.5

Size Categories Used in Regression Analysis

Size Variable	Freestanding Hospitals ADC	Rehabilitation Units ADC
Small	Less than 25	Less than 10
Medium	25-49	10-24
Large	50 or more	25 or more

Time Since Certification

In the interim report, we used the dates indicated in the Online Survey and Certification Reporting system (OSCAR) to develop three categories for certification: before 1985, 1985-1990, and 1991 or later. In contrast to our earlier 1997 work, we did not find that newer facilities were significantly more expensive. However, we did find that there were problems with the certification dates, which frequently differed from the certification dates reported on the HCRIS 12 cost report. (Certification date is not reported on later versions of the cost report.) Upon investigation, we found that the certification date is commonly updated when there is a change in number of beds. We wanted a measure of how long a facility has been operating under TEFRA. Since the change in number of beds does not affect the facility's TEFRA limit, our analysis used the earliest of the following: the certification date reported on the HCRIS 12 cost report, the certification date reported in OSCAR, or the first year for which we have MEDPAR data (1996 or later).

Type of Control

We used a dummy variable to explore whether the type of ownership affects costs. The fully specified regressions in the interim report

indicated that proprietary facilities are more costly than nonprofit or government institutions.

METHODS: MULTIVARIATE REGRESSION ANALYSES, PAYMENT SIMULATIONS, AND SUPPLEMENTAL ANALYSES

We used multivariate regression to examine factors that may explain facility-level variation in costs per case. Our dependent variable was each facility's total (operating and capital) cost per case. We used the natural logarithm to transform cost and examined different specifications and variable measures. We used the coefficients from selected regressions as potential payment adjustments and then simulated payments. Finally, we performed additional analyses to understand the results of our regressions and refined our methodology. Our analyses focused on several issues raised in the interim report, by the TE, or by our subsequent analyses affecting the facility adjustments. We evaluated the following:

- Effects of using case mix measures derived from the compressed or decompressed relative weights.
- Effects of different transformations of the low-income and teaching variables: $\log(1 + \text{ratio})$; $\log(.0001 + \text{ratio})$; and a non-logged form.
- Effects of different measures of commitment to serving low-income patients: the low-income patient percentage, the disproportionate share patient percentage, and the Medicare SSI percentage.
- Sources of cost variation by geographic location and by type of facility.

Table 7.6 provides an overview of the regression models reported in this section.

Table 7.6

Overview of Regression Models Included in This Report

Table	Model No.	Type	Cost per Case Stan-dardized	Explanatory Variables (rural in all models)			
				Compressed or Decompressed CMI	Low-Income Patient Measure	Teaching Measure	Large Urban
7.8	1	F/S	No	Compressed	log(1 + DSH)	log(1 + R2ADC)	Yes
	2	F/S	No	Decompressed	log(1 + DSH)	log(1 + R2ADC)	Yes
7.9	3	Pay	No	Compressed	log(1 + DSH)	log(1 + R2ADC)	Yes
	4	Pay	No	Decompressed	log(1 + DSH)	log(1 + R2ADC)	Yes
7.10	5	Pay	No	Compressed	log(1 + DSH)		Yes
	6	Pay	No	Decompressed	log(1 + DSH)		Yes
7.11	7	Pay	Yes	Compressed	log(1 + DSH)		
	8	Pay	Yes	Decompressed	log(1 + DSH)		
7.12	9	Pay	Yes	Compressed	log(1 + DSH)		
	10	Pay	Yes	Decompressed	log(1 + DSH)		
7.13	10A	Pay	Yes	Decompressed	log(1 + LIP)		
	10B	Pay	Yes	Decompressed	log(1 + SSI)		
7.15	10C	Pay	Yes	Decompressed	log(.0001 + DSH)		
	10D	Pay	Yes	Decompressed	DSH		

F/S = fully specified.

Fully Specified Regressions

We performed evaluation regressions to understand the various
factors affecting cost per case. Our dependent variable was the log of
cost per case. We performed each regression using the log of the CMI
(average CMG weight per case) and log (.705 $*$ WI + .295) on the right-
hand side of the equation. In addition to potential payment variables,
we included dummy variables to indicate those facilities for which we
imputed the low-income patient measure, freestanding units, date of
certification category, size category, type of ownership, and geographic
location.

In the interim report, we examined whether the payment coefficients
were sensitive to whether the payment regression is facility-weighted or
case-weighted. We determined that the case-weighted regressions
performed slightly better. Relying on our earlier analysis, we used
case-weighted regressions for this report.

Payment Regressions

We also performed a set of payment regressions using only those
variables that measure factors used as payment parameters by the
Medicare program. In the first set of payment regressions, we left CMI
and WI on the right side of the equation so that we could measure the

effect of dropping the explanatory variables from the regression. In the second set, we moved CMI to the left-hand side of the equation since payment will be proportional to the CMI. We also moved WI to the left-hand side by standardizing the labor-related share of each facility's CMI-adjusted cost per case for the WI, consistent with the way these adjustments are applied in the IRF PPS. We regressed the CMI and WI-adjusted cost per case on the potential payment variables. We dropped variables that were not significant and repeated the payment regressions using only those variables that were found to be significant. In each set of payment regressions, we compared the effects of using the compressed and decompressed relative weights and of including the large urban variable. In the last set, we also explored the effects of using alternative measures of serving low-income patients and different transformations of those measures.

Payment Simulations

We simulated payments to assess the appropriateness of the adjustments indicated by the payment regressions and other analyses. We used the coefficients from the regressions to determine the payment adjustment factors. For this purpose, we used 1999 claims data from 694 facilities in our analysis file. In each simulation, total payments equal total costs for the cases in the sample. We assumed a 3 percent outlier policy consistent with our recommendation to CMS. We compared payments for selected models relative to a base model that includes a wage adjustment only.

We focused on the payment-to-cost ratios across groups of rehabilitation facilities as an indicator of how well the payment adjustments indicated by the regressions would work to avoid rewarding or penalizing hospitals for sources of cost variation beyond their control. In theory, we wanted to match a facility's payments with the costs it would incur in delivering inpatient rehabilitation services efficiently. However, we did not have a measure of efficiency that allowed us to make this judgment on a facility-by-facility basis. Instead, we assumed that, on average, hospitals operate efficiently, and we used the PTC ratio to assess how well the payment adjustments accounted for cost variations beyond the control of facilities. In doing

so, we assumed that inefficiency is not correlated with certain hospital characteristics such as geographic location and commitment to serving low-income patients and that cost variation across groups is not attributable to inefficiencies. We know, however, that the TEFRA system has forced older facilities to be more efficient than newer facilities and that this might bias the PTC ratios for certain groups of facilities. Therefore, in addition to examining the PTC ratios, we were interested in exploring the characteristics within each class of facilities that would receive substantially more than their costs under alternative payment policies and those that would receive substantially less.

Supplemental Analyses

We performed additional analyses to enhance our understanding of the regression and simulation results. For example, we performed additional regressions and descriptive analyses directed at specific issues such as the choice of the low-income patient measure and the characteristics of specific classes of hospitals.

We also used the added variable plot technique to identify hospitals that unduly influence the regression results for specific variables. Using the wage index variable as an example, we

- regressed cost per case on all independent variables other than the wage index and saved the studentized residuals from this model
- regressed the wage index variable on all other independent variables in the model and saved the studentized residuals from this model
- plotted the residuals from the first regression against those from the second regression
- superimposed a line onto this plot that has an intercept of 0 and a slope equal to the coefficient for the wage index.

A facility could be influential on the regression if either its wage index value is extreme or it is a large facility. While one cannot entirely separate the effects of facility size from undue influence, the plotted points are made proportional to the number of cases per facility. The proportional plotted points help to account for the size of facilities that deviate from the superimposed line and are thus not well explained by the regression coefficient for the wage index.

ANALYSIS RESULTS

Characteristics of Facilities in Analysis File

Table 7.7 summarizes key characteristics of the 714 facilities in
our analysis file. There are 90 more facilities and 28 percent more cases
than there were in our analysis file for the interim report. The case-
weighted CMI is comparable across urban and rural locations using both the
compressed and decompressed weights. Consistent with the interim report,
rural facilities tend to have fewer cases and a higher average cost per
case.

Eighty-nine percent of the facilities for which we have actual data
on SSI and Medicaid ratios have a low-income patient ratio below .20.
There are 38 facilities with a low-income patient ratio between .20 and
.30 and 19 facilities with a low-income patient ratio of .30 or more.
Compared to the interim report, there are relatively fewer facilities with
a high low-income patient ratio. This is probably attributable to the
change in the way we impute missing Medicaid patient percentages. In the
interim report, if we did not have information on the number of Medicaid
patients in a rehabilitation unit, we used the Medicaid percentage for the
acute care portion of the facility (which tends to be higher). For this
report, we treated these facilities as having missing values.

There are 104 facilities with teaching programs in our analysis
file, five fewer than in the interim report. Only 40 facilities (including
five freestanding hospitals) have a resident-to-ADC ratio equal to or
greater than .10. The teaching facilities tend to be located in large
urban areas and have a higher case mix and a higher proportion of low-
income patients.

Table 7.7

Characteristics of Facilities in Analysis File

				Case-Weighted						Facility-Weighted			
	No. of Facil- ities	Avg. Annual Cases	Cost per Case ($)	CMI Com- pressed	CMI Decom- pressed	Capital As % of Cost	DSH Ratio	Resi- dent to ADC	Wage Index	% Large Urban	% Rural	% Free- stand- ing	ADC
All facilities	714	320	12,073	0.999	0.999	10	0.121	0.014	0.973	46	11	24	25
By geographic area													
Urban	633	336	12,035	0.999	0.999	10	0.122	0.015	0.983	52	0	25	27
Large urban	329	319	12,404	0.995	0.995	10	0.125	0.022	1.052	100	0	21	28
Other urban	304	354	11,674	1.003	1.003	11	0.119	0.009	0.916	0	0	29	26
Rural	81	196	12,590	0.994	0.995	11	0.111	0.001	0.828	0	100	15	13
By region													
New England	27	422	13,737	1.002	1.001	9	0.104	0.002	1.106	63	4	41	36
Middle Atlantic	91	473	11,577	0.917	0.909	9	0.103	0.027	1.083	63	5	27	40
South Atlantic	108	417	11,215	1.004	1.005	11	0.142	0.008	0.932	44	11	28	31
East North Central	153	254	12,214	1.011	1.012	8	0.097	0.021	0.976	41	12	8	18
East South Central	39	503	11,530	1.054	1.059	11	0.154	0.016	0.860	28	18	36	35
West North Central	63	214	11,655	1.007	1.008	10	0.098	0.012	0.891	35	19	13	16
West South Central	122	303	12,507	0.997	0.997	13	0.128	0.006	0.888	46	13	36	24
Mountain	38	231	10,838	0.952	0.949	12	0.106	0.006	0.952	32	11	32	18
Pacific	73	154	16,230	1.183	1.200	11	0.173	0.011	1.153	60	7	16	15
Average daily census													
Units of acute care hospitals	545	245	12,000	1.001	1.001	9	0.120	0.018	0.980	48	13	0	18
Less than 10	168	117	12,591	0.987	0.987	10	0.090	0.005	0.936	38	27	0	7
10-24	274	234	11,698	0.993	0.992	9	0.114	0.015	0.969	50	8	0	16
25 or more	103	482	12,158	1.017	1.018	9	0.139	0.027	1.011	57	1	0	39
Freestanding	169	562	12,176	0.996	0.996	12	0.124	0.008	0.963	41	7	100	49
Less than 25	32	185	15,889	1.082	1.090	12	0.154	0.000	0.968	38	16	100	16
25-49	73	462	12,428	1.000	1.000	12	0.120	0.004	0.921	36	5	100	38
50 or more	64	864	11,624	0.985	0.983	12	0.122	0.012	0.987	48	5	100	77

- 250 -

Table 7.7 (cont.)

	No. of Facilities	Avg. Annual Cases	Cost per Case ($)	CMI Compressed	CMI Decompressed	Capital As % of Cost	DSH Ratio	Resident to ADC	Wage Index	% Large Urban	% Rural	% Free-standing	ADC
			Case-Weighted							Facility-Weighted			
Low-income patient ratio													
< .10	325	331	11,233	0.962	0.958	10	0.071	0.006	0.988	46	12	22	24
.10-.19	154	339	13,223	1.031	1.034	10	0.162	0.025	0.998	48	12	30	29
.20-.29	38	296	14,439	1.088	1.096	9	0.285	0.057	0.974	55	8	37	30
30% or more	19	231	15,787	1.176	1.193	10	0.465	0.050	1.048	63	0	32	30
Missing	178	299	11,837	1.009	1.010	11	0.119	0.007	0.910	41	12	17	21
Teaching status													
No teaching	610	308	11,854	1.000	1.000	11	0.115	0.000	0.949	42	13	24	23
Resident-to-ADC ratio													
< .1	64	438	12,715	0.966	0.963	9	0.121	0.034	1.089	70	6	23	41
.1-.2	24	411	13,743	1.057	1.062	8	0.234	0.144	1.093	63	0	17	43
=> .2	16	171	14,539	1.039	1.042	9	0.188	0.309	0.985	69	0	6	20
Type of ownership													
Voluntary	451	291	11,913	1.000	1.000	9	0.118	0.019	0.993	47	11	11	22
Proprietary	200	416	12,044	0.990	0.990	14	0.120	0.003	0.942	46	10	57	34
Government	63	225	13,735	1.037	1.039	8	0.164	0.031	0.967	38	21	10	21
Certification date													
Before 1/1/1985	224	350	11,955	1.007	1.008	9	0.128	0.018	1.002	50	5	20	29
From 1/1/85 to 12/31/90	252	329	11,713	0.994	0.993	11	0.111	0.012	0.957	46	12	24	25
After 1/1/1991	234	285	12,652	0.996	0.996	11	0.127	0.012	0.959	42	16	27	21
Missing	4	126	13,167	1.014	1.016	10	0.053	0.000	0.866	25	50	0	8
Medicare days as % of inpatient days													
0-49%	129	187	14,352	1.078	1.086	9	0.180	0.047	1.018	65	4	19	24
50-64%	168	348	12,710	1.006	1.006	10	0.139	0.021	1.031	52	4	21	30
65-79%	252	363	11,473	0.987	0.986	11	0.105	0.008	0.963	44	14	28	25
80% and over	163	333	11,367	0.977	0.974	11	0.103	0.002	0.906	28	21	23	20
Missing	2	41	24,996	1.262	1.289		0.121	0.000	0.922	100	0	0	2
Alaska/Hawaii	2	408	11,640	1.030	1.031	6	0.154	0.000	1.199	0	0	50	39
Total costs per case													
< 25th percentile	178	389	8,393	0.926	0.920	10	0.089	0.008	0.940	47	9	17	23

Table 7.7 (cont.)

			Case-Weighted							Facility-Weighted			
	No. of Facil- ities	Avg. Annual Cases	Cost per Case ($)	CMI Com- pressed	CMI Decom- pressed	Capital As % of Cost	DSH Ratio	Resi- dent to ADC	Wage Index	% Large Urban	% Rural	% Free- stand- ing	ADC
26-50th percentile	179	369	11,002	0.983	0.981	11	0.123	0.007	0.946	40	16	28	28
51-75th percentile	178	301	13,523	1.035	1.038	11	0.136	0.025	0.993	46	11	22	26
> 75th percentile	179	221	18,359	1.105	1.114	11	0.156	0.021	1.047	52	9	27	23
Case mix index													
< 25th percentile	178	362	10,114	0.858	0.845	10	0.086	0.016	1.011	47	14	18	24
26-50th percentile	179	377	11,433	0.972	0.969	11	0.104	0.008	0.955	44	12	31	27
51-75th percentile	178	352	12,843	1.057	1.062	11	0.148	0.013	0.944	44	13	30	28
> 75th percentile	179	190	15,648	1.214	1.234	9	0.173	0.025	0.988	49	6	16	20
Wage index													
< 25th percentile	177	322	11,577	1.011	1.012	11	0.128	.004	0.829	7	36	27	23
26-50th percentile	184	313	11,567	1.003	1.004	11	0.128	0.012	0.911	38	7	23	24
51-75th percentile	187	325	11,379	0.981	0.980	10	0.111	0.018	0.983	68	1	24	25
> 75th percentile	166	321	13,944	1.002	1.001	9	0.124	0.023	1.181	72	2	21	29

Fully Specified Regressions

Table 7.8 displays the results of two fully specified regressions. Model 1 uses a CMI derived from the compressed relative weights and Model 2 uses a CMI derived from the decompressed relative weights. The other explanatory variables used in the regressions are identical. The CMI in both models is highly significant. The CMI coefficient is consistent with the relative weights: 1.255 using the compressed weights (Model 1) and 1.146 using the decompressed weights (Model 2). The R-squared is slightly higher using the compressed weights (Model 1: .4659 versus Model 2: .4649).

The coefficients and significance of the remaining variables in Models 1 and 2 are quite similar. The WI is highly significant. DSH is also significant. Cost per case increases about 3.5 percent for each .10 increment in the DSH ratio. The dummy variable for location in a rural area is positive and significant while the large urban area variable is negative and significant. Facilities located in rural areas are about 15

Table 7.8

Fully Specified Regression Results Comparing Compressed and Decompressed CMIs

| | Model 1 | | Model 2 | |
| | Compressed CMI | | Decompressed CMI | |
Variable	Coef.	t-stat.	Coef.	t-stat.
Log CMI	1.255	17.830	1.146	17.777
Log WI	1.247	12.062	1.252	12.101
Log (1 + DSH ratio)	0.357	2.805	0.359	2.817
DSH gap-fill dummy	0.029	1.452	0.029	1.443
Log (1 + residents to ADC)	0.340	1.645	0.342	1.651
Freestanding	0.053	2.571	0.053	2.540
Small size	0.081	2.882	0.080	2.875
Large size	-0.029	-1.569	-0.029	-1.574
Certification before 1985	-0.018	-.904	-0.018	-.893
Certification after 1990	0.048	2.323	0.048	2.332
Large urban area	-0.050	-2.590	-0.050	-2.588
Rural	0.141	4.051	0.140	4.032
Proprietary	0.029	1.354	0.029	1.340
Intercept	9.322	374.910	9.323	374.326
R-squared	0.4659		0.4649	

percent more costly than urban facilities; large urban facilities are about 5 percent less costly than facilities located in other urban or rural areas after controlling for other factors affecting costs.[13] Teaching is not significant.[14]

Among the remaining explanatory variables used in the fully specified regressions reported in Table 7.8, the dummy variables for type of facility, new facilities, and size are significant in both Models 1 and 2. Freestanding hospitals are about 5 percent more costly than rehabilitation units of acute care hospitals, and small facilities are about 8 percent more costly than medium-sized ones after controlling for other factors affecting costs. Facilities certified after 1990 are about 5 percent more costly than older facilities. Type of ownership, which was a significant explanatory variable in the interim report, is not significant using these models.

An area of concern is that the coefficient for WI is statistically different from the expected value of 1.0.[15] This result, which is different from our interim report finding, implies that the wage index may overstate the resources required in low-wage areas and understate the resources required in high-wage areas.

Payment Regressions

In the payment regressions, we retained only those explanatory variables that were potential payment factors. We eliminated the variables for type of facility, size, date of certification, and type of ownership. Model 3 uses the compressed CMI and Model 4 uses the decompressed CMI. The

---

[13] Based on e raised to the power of the coefficient for the variable where e equals the natural anti-log of 1.0. For example, the rural differential indicated by the coefficient for rural areas in Model 1 is exp(0.141) = 1.1514.

[14] The regression results are similar when we use alternative low-income patient measures and different forms of the low-income and teaching variables. Although we present the results using the resident-to-ADC ratio as our measure of teaching intensity, we evaluated three other measures: the resident-to-bed ratio, a dummy variable for teaching, and a dummy variable for major participating facilities. None was significant.

[15] The wage index coefficient for Model 1 (1.247) has a standard error of .104. The wage index coefficient for Model 2 (1.252) has a standard error of .103.

R-squareds drop slightly from about .465 in the fully specified regressions (Models 1 and 2 in Table 7.8) to .43 (Models 3 and 4 in Table 7.9). Large urban continues to be significant and negative, and teaching remains insignificant. The wage index variable in Models 3 and 4 (Table 7.9) is no longer statistically distinguishable from the expected 1.0 value.

In the regressions reported in Table 7.10, we dropped the teaching variable because it is not significant. Model 5 uses the compressed CMI and Model 6 uses the decompressed CMI. Without the teaching variable, there is a slight increase in the coefficient for the DSH variable. Otherwise, the coefficients for the remaining variables are nearly the same as those reported in Table 7.9 for models that include the teaching variable.

Table 7.9

Payment Regressions Comparing Compressed and Decompressed CMIs, Using Cost per Case as Dependent Variable with CMI, WI, DSH, Residents-to-ADC, Rural, and Large Urban as Explanatory Variables

Variable	Model 3		Model 4	
	Compressed CMI		Decompressed CMI	
	Coef.	t-stat.	Coef.	t-stat.
Log CMI	1.246	17.275	1.139	17.235
Log WI	1.177	11.261	1.183	11.307
Log (1 + DSH ratio)	0.366	2.819	0.367	2.827
DSH gap-fill dummy	0.024	1.187	0.024	1.184
Log (1 + residents to ADC)	0.148	0.706	0.150	0.718
Large urban area	-0.052	-2.652	-0.052	-2.651
Rural	0.171	4.956	0.171	4.941
Intercept	9.358	463.048	9.359	463.346
R-squared	0.4308		0.4300	

Table 7.10

**Payment Regressions Comparing Compressed and Decompressed CMIs Using Cost
per Case as Dependent Variable with CMI, WI, DSH, Rural,
and Large Urban as Explanatory Variables**

Variable	Model 5		Model 6	
	Compressed CMI		Decompressed CMI	
	Coef.	t-stat.	Coef.	t-stat.
Log CMI	1.245	17.270	1.138	17.230
Log WI	1.186	11.454	1.192	11.502
Log (1 + DSH ratio)	0.389	3.095	0.391	3.108
DSH gap-fill dummy	0.024	1.158	0.024	1.154
Large urban area	-0.052	-2.623	-0.052	-2.621
Rural	0.171	4.948	0.171	4.932
Intercept	9.357	463.523	9.358	462.818
R-squared	0.4304		0.4296	

Next, we standardized cost per case for the CMI and WI; we report the results in Table 7.11. We standardized by the compressed CMI and the decompressed CMI in Models 7 and 8, respectively. When the dependent variable is the CMI and wage-adjusted cost per case, the R-squared drops to .07 in both models. The lower R-squared value is primarily because we accounted for the variables with the greatest explanatory powers (CMI and WI) on the left side of the equation. The large urban variable remains slightly significant and negative. As explained in the subsection below on supplemental analyses, we undertook additional analyses to investigate the large urban effect on standardized costs per case. On the basis of these analyses, we decided to drop large urban as an independent variable in our remaining payment regressions.

Table 7.11

**Payer Regressions Comparing Compressed and Decompressed CMIs Using CMI-
and WI-adjusted Cost per Case as Dependent Variable with DSH, Rural,
and LU as Explanatory Variables**

| | Model 7 | | Model 8 | |
| | Compressed | | Decompressed | |
Variable	Coef.	t-stat.	Coef.	t-stat.
Log (1 + DSH ratio)	0.552	4.719	0.490	4.210
DSH gap-fill dummy	0.019	0.955	0.017	0.873
Large urban area	-0.037	-2.133	-0.036	-2.060
Rural	0.158	4.626	0.158	4.630
Intercept	9.328	500.295	9.336	502.778
R-squared	0.0718		0.0657	

Table 7.12 shows a comparison between the regression results using
the compressed (Model 9) and decompressed (Model 10) weights, using the
CMI- and WI-adjusted cost per case as the dependent variable and only the
DSH variables and rural area dummy variable as the explanatory variables.
The R-squared is slightly higher in Model 9, which uses the compressed CMI
to standardize cost per case, than in Model 10, which uses the
decompressed CMI. The coefficient for the DSH variable is slightly higher
in Model 9. In both models, the dummy variable for imputed DSH ratios
remains insignificant. This confirms that the changes in our imputation
method led to an improvement over the method in our interim report.

Table 7.12

**Payer Regressions Comparing Compressed and Decompressed CMIs Using
CMI- and WI-adjusted Cost per Case as Dependent Variable with
DSH and Rural as Explanatory Variables**

| | Model 9 | | Model 10 | |
| | Compressed | | Decompressed | |
Variable	Coef.	t-stat.	Coef.	t-stat.
Log (1 + DSH ratio)	0.545	4.651	0.484	4.146
DSH gap-fill dummy	0.023	1.176	0.0216	1.087
Rural	0.177	5.317	0.175	5.303
Intercept	9.309	564.478	9.317	567.447
R-squared	0.0658		0.0601	

We used Model 10 as our base model in additional payer regressions
to examine alternative measures of commitment to serving low-income
patients and different transformations of the low-income variable.

Measure of Low-Income Population

Table 7.13 compares the effects of using DSH (Model 10), LIP (Model
10A), and the Medicare SSI ratio (Model 10B) as the measure of commitment
to serving low-income patients. The dependent variable in all three models
is cost per case standardized by decompressed CMIs and the WI. We analyzed
a measure using the SSI ratio only because of concerns over the
reliability of the Medicaid utilization data. The regressions using DSH
(Model 10) and LIP (Model 10A) have the same R-squared value (.060), while
the R-squared value is lower (.0517) when SSI is used. The similarity of
the explanatory power of the LIP and DSH variables is consistent with the
interim report results. Either variable, regardless of concerns over the
Medicaid data, performs better than the SSI variable alone.

Additional information in the UDSmr FIM data allowed us to explore
the effect of alternative measures of serving a low-income patient
population. The purpose was to determine whether the alternatives
provide greater explanatory power than the measures that are based on
currently available data for rehabilitation facilities, and therefore
might be considered in future refinements of the low-income patient

Table 7.13

**Payer Regressions Comparing Different Low-Income Patient Measures Using
Decompressed CMI- and WI-adjusted Cost per Case as Dependent Variable with
Low-Income Patient Measure and Rural as Explanatory Variables**

	Model 10		Model 10A		Model 10B	
	Log (1 + DSH ratio)		Log (1 + LIP ratio)		Log (1 + SSI ratio)	
Variable	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.
Log (1 + low-income measure)	0.484	4.146	0.589	4.168	0.556	3.442
Low-income patient gap-fill	0.0216	1.087	0.023	1.163	---	---
Rural	0.175	5.303	0.172	5.225	0.169	5.091
Intercept	9.317	567.447	9.317	563.471	9.34	679.162
R-squared	0.0601		0.0603		0.0517	

adjustment. We have UDSmr FIM data for 605 facilities in our analysis file. (Our HealthSouth data do not include information on the non-Medicare population and cannot be used for this analysis.) Table 7.14 compares regression results using DSH, LIP, the SSI ratio for Medicare beneficiaries, the dual eligible (DUAL) ratio (proportion of Medicare patients who are also entitled to Medicaid), and an all-low-income patient (ALL-LIP) ratio that is defined as the sum of the dual eligible, Medicaid, self-pay, and charity care days as a percentage of total inpatient days. (The DUAL and ALL-LIP measures cannot be established from currently available data for all rehabilitation facilities.)
We also examined a measure based on the proportion of non-Medicare patients who are entitled to Medicaid. We do not present the results for this regression because the coefficient for the ratio was not significant. In each regression, the dependent variable is the log of the cost per case standardized by the decompressed CMI and the wage index. The low-income measure is expressed in the log (1 + ratio) form.

Since the UDSmr facilities are a subset of the facilities in the analysis file, the coefficients for DSH and LIP differ from those resulting from the regressions on the full analysis file of 714 facilities. We found that the ALL-LIP variable does not add any explanatory power relative to the DSH or LIP variables for this set of facilities. In fact, DSH has a higher R-squared than either LIP or ALL-LIP. The R-squared for DUAL (.0496) is the same as for LIP and higher

Table 7.14

Payer Regressions Comparing Alternative Low-Income Patient Measures Using UDSmr Facilities with Decompressed CMI- and WI-adjusted Cost per Case as Dependent Variable and Low-Income Patient Measure and Rural as Explanatory Variables

	DSH		LIP		SSI		DUAL		ALL-LIP	
Variable	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.
Log (1 + low-income ratio)	0.476	3.687	0.537	3.462	0.545	3.169	0.434	3.443	0.444	3.662
Rural area	0.170	4.412	0.168	4.349	0.162	4.366	0.163	4.223	0.164	4.258
Intercept	9.328	546.804	9.33	546.011	9.343	632.02	9.34	594.107	9.324	519.637
R-squared	0.0520		0.0495		0.0462		0.0496		0.0517	

than for SSI (.0462). Consistent with our regression results in Table 7.13
for the full analysis file, the SSI variable has the lowest explanatory
power. Differences in the size of the coefficients for the low-income
patient measures are in the expected directions. Those measures that have
the higher ratios have lower coefficients (e.g., DUAL relative to SSI
(.434 versus .545); ALL-LIP relative to LIP (.444 versus .537)).

Transformation of the Low-Income Variable

Table 7.15 identifies the effects of using different transformations
of the low-income variable: log (1 + DSH ratio) (Model 10), log (.0001 +
DSH ratio) (Model 10C), and in a non-logged form (Model 10D). The
dependent variable in the regressions is cost per case standardized by the
decompressed CMIs and WI. Rural and a dummy for gap-filled low-income
patient values are included as explanatory variables. Consistent with the
interim report, the R-squareds are higher using log (.0001 + DSH ratio)
for the low-income variable (.0652) than they are using log (1 + DSH
ratio) (.0601) or the non-logged DSH ratio (.0583). The coefficients for
the DSH variable are different because the scales are different. The rural
coefficient is not affected by how the low-income variable is formulated.

Table 7.15

**Payer Regressions Comparing Different Transformations of the DSH Patient
Measures Using Decompressed CMI- and WI-adjusted Cost per Case as
Dependent Variable with DSH and Rural as Explanatory Variables**

	Model 10		Model 10C		Model 10D	
	Log (1+ DSH ratio)		Log (.0001 + DSH ratio)		Non-logged DSH ratio	
Variable	Coef.	t-stat.	Coef.	t-stat.	Coef	t-stat.
DSH variable	0.484	4.146	0.055	4.604	0.385	3.971
DSH gap-fill dummy	0.0216	1.087	0.018	0.930	0.220	1.104
Rural	0.175	5.303	0.175	5.300	0.175	5.300
Intercept	9.317	567.447	9.500	322.592	9.325	603.631
R-squared	0.0601		0.0652		0.0583	

The logged and non-logged transformations of the DSH ratio affect
the low-income patient adjustment for different patient percentages. To
illustrate, Table 7.16 compares the adjustments using the coefficients
from Models 10, 10C, and 10D at different DSH ratios. We used the range
from 0.00 to 0.75 because it covers up to the 99th percentile of the DSH

ratios in our analysis sample. Ninety-five percent of cases are in facilities with a DSH ratio of .30 or less. In our preliminary analyses for the interim report, CMS requested that the low-income patient adjustment be expressed as an add-on to a standard payment amount. Consistent with this request, we show the adjustment factors resulting from the regression coefficients after normalizing by the adjustment factor that would apply for a DSH ratio equal to 0.

In the interim report, we recommended basing the low-income patient adjustment on the log (.0001 + ratio) form. Although this form continues to have the highest explanatory power, it also has the disadvantage of sharply increasing payments to facilities with relative low DSH ratios. As seen in Table 7.16, a single percentage point increase in a very low DSH ratio can result in a substantial increase in the low-income patient adjustment when the log (.0001 + DSH ratio) form is used to establish the payment adjustment. An increase from .02 to .03 in the DSH ratio increases payment by more than two percentage points. As the DSH ratio increases, however, the incremental increase in payment flattens out. For example, an increase from .30 to .35 in the DSH ratio increases payment by less than one percentage point.

Figure 7.1 compares the slope of the DSH adjustments using the three forms of the DSH variable. To facilitate this comparison, we normalized the adjustments to the adjustment factor for a facility with a DSH ratio of .121, which is the case-weighted average DSH ratio in our analysis file (including the gap-filled values). As the figure illustrates, the slopes of the three forms are similar for the great majority of hospitals that are found in the mid-range of the DSH ratios. Compared to the log (1 ± DSH) form (Model 10), the non-logged form (Model 10D)

---

[16] We used the following formula for the low-income patient adjustment based on the regression coefficients from Model 10C: ((.0001 + DSH) raised to the power of .0547)/(.0001 raised to the power of .0547). Normalization to 0 DSH patient percentage results in a considerably lower standard payment rate than the average payment to a hospital with a typical DSH patient percentage. Normalization is not needed for the other forms.

Table 7.16

Comparison of Case-Weighted Low-Income Patient Adjustment Factors
Using Different Forms of the DSH Variable Normalized to
Adjustment Factor for 0 DSH

DSH Patient Percentage	No. of Providers	No. of Cases	Normalized to 0 Adjustment Factor		
			Model 10	Model 10C	Model 10D
			$(1 + DSH)$ $**.4838$	$(.0001+DSH)**.0547/$ $(.0001**.0547)$	$e **(DSH*.3855)$
0	1	72	1.0000	1.0000	1.0000
>0=<.01	9	869	1.0038	1.2700	1.0030
>.01=<.02	15	4,316	1.0079	1.3218	1.0063
>.02=<.03	32	10,808	1.0125	1.3558	1.0101
>.03=<.04	31	8,702	1.0169	1.3784	1.0137
>.04=<.05	26	9,881	1.0212	1.3958	1.0172
>.05=<.1	226	80,343	1.0360	1.4373	1.0297
>.1=<.15	152	52,782	1.0576	1.4756	1.0297
>.15=<.2	93	28,911	1.0787	1.5020	1.0676
>.2=<.25	50	14,392	1.1019	1.5243	1.0894
>.25=<.3	29	8,887	1.1233	1.5411	1.1104
>.3=>.35	12	2,646	1.1474	1.5573	1.1351
>.35=<.4	12	2,441	1.1681	1.5694	1.1571
>.4=<.45	6	1,123	1.1858	1.5788	1.1767
>.45=<.5	2	146	1.2039	1.5876	1.1975
>.5=<.55	2	409	1.2355	1.6015	1.2354
>.55=<.6	2	159	1.2426	1.6044	1.2441
>.6=<.65	1	309	1.2707	1.6152	1.2802
>.65=<.7	2	429	1.2817	1.6192	1.2948
>.7=<.75	2	908	1.3050	1.6272	1.3269
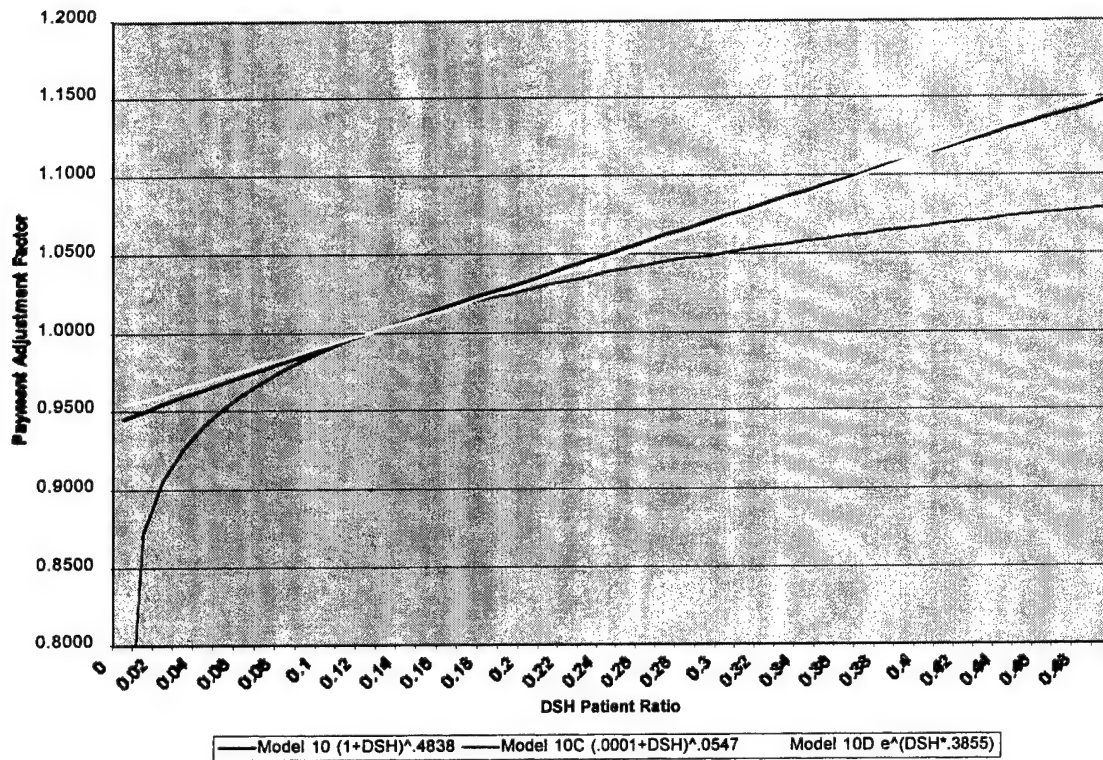
Figure 7.1--Comparison of Low-Income Patient Adjustment Factors
Resulting from Different Forms of the DSH Variable
Normalized for Average DSH Ratio (12.1 percent)

has a slightly flatter slope. It provides a slightly higher DSH adjustment
for relatively low DSH patient ratios and a slightly lower DSH adjustment
for facilities with relatively high DSH patient ratios. Both provide
higher payments than the log (.0001 + DSH ratio) form (Model 10C) at
relatively low and high DSH ratios. A minimum threshold for receiving an
adjustment for serving low-income patients, e.g., a DSH ratio of at least
.05, would avoid the effects of the steep slope at the low DSH ratios
using the log (.0001 + DSH ratio) form (Model 10C).

Payment Simulations

We used payment simulations to confirm the results of the regression analyses. We individually priced the 1999 cases in our analysis file assuming a 3 percent outlier policy and budget neutrality (total payments equal total costs for the cases). There are 694 facilities with 247,461 cases in the simulations.[17] We used the coefficients from the regressions to establish the payment parameters. We normalized the low-income patient adjustment when applicable so that the adjustment for 0 low-income patients was 1.0.

For the most part, we report the payment simulations using the decompressed weights consistent with the CMS decision to use the decompressed relative weights in the IRF PPS. We used Model 10 as our reference model. It uses the decompressed relative weights, 1 + DSH ratio as the low-income patient measure, and an adjustment for rural facilities. For Model 10C, which uses the log (.0001 + DSH ratio) form with a rural adjustment, we also evaluated the effects of a 5 percent threshold (Model 10CT). No DSH payments would be made to facilities with a DSH ratio of less than .05. The payment adjustment factor would be normalized to the adjustment for the .05 DSH ratio to avoid a notch effect. Since the pattern of PTC ratios resulting from changing the payment parameters is similar when the compressed relative weights are used, we report only the results for Model 9. For comparison purposes, we also include a preliminary simulation using only the wage index as a facility-level adjustment (base model).

We report the payment simulation results for the following facility-level adjustments:[18]

Base model: (0.295 + 0.705 * WI)

Model 9: (0.295 + 0.705 * WI) * (1 + 0.1928 * rural) * ((1.0 + DSH) ** 0.5451)

---

[17] The numbers of cases and facilities in the simulations differ somewhat from those in the analysis file used for the regressions. The payment simulations use 1999 claims, whereas our regression analyses are based on pooled data for 1998-1999.

[18] The exponentiated form of the rural adjustment is shown in the formula (e.g., e ** 0.175 = 1.1914).

Model 10: (0.295 + 0.705 * WI) * (1 + 0.1914 * rural) * ((1.0 +

    DSH) ** 0.4838);

Model 10A: (0.295 + 0.705 * WI) * (1 + 0.1882 * rural) * ((1 +

    LIP) ** 0.5893)

Model 10C: (0.295 + 0.705 * WI) * (1 + 1907 * rural) *

    ((.0001 + DSH) ** 0.0547)/(.0001 ** 0.0547)

Model 10CT:

    IF DSH GT 0.050: (0.295 + 0.705 * WI) * (1 + .1907 * rural) *

        ((.0001 + DSH) ** 0.0547)/(.0501 ** 0.0547)

    ELSE: (0.295 + 0.705 * WI) * (1 + 0.1907 * rural)

We did not simulate Models 10B (log (1 + SSI ratio)) and 10D (non-logged DSH ratio) since they do not provide as good a fit with the data as the other alternatives. Table 7.17 summarizes the PTC ratios resulting from the simulations by key facility characteristics. More-detailed simulation results are available from the authors.

Base Model with WI Adjustment Only

The base model simulation adjusts for the payment amount for the CMG to which the patient is assigned using the decompressed relative weights and for the wage index. Table 7.17 shows that without additional facility-level adjustments, large urban hospitals would have a slightly higher than average PTC ratio of 1.02, whereas rural facilities would have, on average, a PTC ratio of .87. Facilities with low-income patient ratios of less than .10 would have an average PTC ratio of 1.03, while those with low-income patient ratios of .10 or greater would have PTC ratios of less than .97. Teaching hospitals with a resident-to-ADC ratio of .20 or greater would have a PTC ratio of .89, while the ratio for other teaching hospitals would be only slightly less than 1.0. There are marked regional differences in the payment-to-cost ratios. The average PTC ratio is .97 or lower for facilities located in the New England (.97), West North Central (.97), and West South Central (.93) regions. The PTC ratios are 1.04 for facilities located in the South Atlantic and Mountain regions. Facilities in the lowest quartile of CMI values would have a PTC ratio of 1.03, while the remaining facilities would have an average PTC ratio of slightly less than 1.0. Units of acute care hospitals would have an average PTC ratio of 1.01 compared to an average ratio of .98 for freestanding facilities.

Table 7.17

Summary Payment Simulation: Payment-to-Cost Ratio

	No. of Cases	No. of Facil.	Cost per Case ($)	Base Model	Model 9	Model 10	Model 10A	Model 10C	Model 10CT
All facilities	247,461	694	11,386	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
By geographic area									
Large urban	112,270	319	11,757	1.0203	1.0128	1.0120	1.0111	1.0095	1.0109
Other urban	118,014	297	10,974	1.0000	0.9881	0.9889	0.9899	0.9913	0.9900
Rural	17,177	78	11,793	0.8678	0.9928	0.9925	0.9925	0.9939	0.9932
By region									
New England	12,015	26	13,058	0.9746	0.9606	0.9605	0.9620	0.9657	0.9640
Middle Atlantic	46,433	90	11,057	1.0230	1.0163	1.0090	1.0100	1.0095	1.0106
South Atlantic	48,420	106	10,667	1.0373	1.0532	1.0524	1.0529	1.0510	1.0503
East North Central	42,199	148	11,280	1.0050	0.9927	0.9942	0.9955	0.9889	0.9930
East South Central	21,029	37	11,077	1.0035	1.0175	1.0216	1.0221	1.0271	1.0237
West North Central	14,896	60	10,981	0.9728	0.9687	0.9711	0.9722	0.9700	0.9718
West South Central	39,885	118	11,707	0.9281	0.9272	0.9277	0.9257	0.9313	0.9290
Mountain	9,988	38	10,262	1.0429	1.0365	1.0344	1.0326	1.0339	1.0358
Pacific	12,596	71	14,994	1.0179	1.0268	1.0378	1.0345	1.0350	1.0331
Unit of acute hospital	142,504	529	11,265	1.0124	1.0138	1.0140	1.0138	1.0133	1.0139
ADC less than 10 pts.	21,281	164	11,819	0.9303	0.9457	0.9469	0.9469	0.9424	0.9477
ADC 10-24 pts.	67,387	262	10,955	1.0159	1.0169	1.0166	1.0164	1.0154	1.0158
ADC 25 or more pts.	53,836	103	11,433	1.0416	1.0379	1.0383	1.0380	1.0396	1.0387
Freestanding hospital	104,957	165	11,551	0.9837	0.9817	0.9814	0.9818	0.9824	0.9816
ADC less than 25	6,541	29	14,711	0.8735	0.8828	0.8884	0.8874	0.8858	0.8859
ADC 25-49	38,265	73	11,738	0.9490	0.9479	0.9486	0.9491	0.9514	0.9503
ADC 50 or higher	60,151	63	11,088	1.0229	1.0187	1.0170	1.0175	1.0172	1.0166
Low-income patient ratio									
Less than .10	150,256	416	10,635	1.0272	1.0075	1.0071	1.0067	1.0059	1.0084
.10-.19	77,622	213	12,255	0.9676	0.9819	0.9826	0.9828	0.9897	0.9864
.20-.29	14,453	46	13,452	0.9484	0.9980	0.9987	1.0003	0.9913	0.9881
.30 or higher	5,130	19	14,430	0.9657	1.0764	1.0741	1.0768	1.0283	1.0248

Table 7.17 (cont.)

	No. of Cases	No. of Facil.	Cost per Case ($)	Base Model	Model 9	Model 10	Model 10A	Model 10C	Model 10CT
Resident-to-ADC ratio									
Non-teaching	204,191	593	11,156	1.0036	1.0019	1.0025	1.0026	1.0024	1.0030
Less than .10	29,679	61	12,166	0.9926	0.9897	0.9863	0.9857	0.9893	0.9875
.10-.19	10,672	24	12,971	0.9929	1.0246	1.0236	1.0250	1.0175	1.0141
.20 or higher	2,919	16	13,722	0.8878	0.8986	0.8984	0.8973	0.8956	0.8953
Medicare days as %									
Less than 50	26,733	126	13,456	0.9459	0.9568	0.9597	0.9540	0.9591	0.9572
50-64	63,268	163	12,028	0.9967	0.9968	0.9961	0.9953	0.9977	0.9956
65-79	98,543	246	10,808	1.0296	1.0242	1.0241	1.0249	1.0236	1.0247
80 and over	58,821	157	10,707	0.9854	0.9881	0.9875	0.9904	0.9867	0.9885
Missing	96	2	21,544	0.8332	0.8206	0.8296	0.8299	0.8332	0.8319
Alaska/Hawaii	781	2	10,383	1.2412	1.2462	1.2415	1.2260	1.2520	1.2473
CMI quartiles									
Q1	71,020	177	9,482	1.0367	1.0403	1.0220	1.0225	1.0215	1.0233
Q2	69,477	171	11,179	0.9959	0.9878	0.9878	0.9883	0.9882	0.9895
Q3	67,239	175	12,018	0.9799	0.9900	0.9901	0.9902	0.9911	0.9900
Q4	39,725	171	14,083	0.9905	0.9862	1.0047	1.0035	1.0033	1.0011
WI quartiles									
Q1	61,896	171	10,898	0.9530	0.9768	0.9780	0.9786	0.9808	0.9793
Q2	59,594	172	10,730	1.0144	1.0100	1.0107	1.0112	1.0126	1.0120
Q3	67,734	188	10,887	1.0288	1.0172	1.0162	1.0150	1.0139	1.0148
Q4	58,237	163	13,156	1.0016	0.9954	0.9949	0.9952	0.9930	0.9940

Comparison of Compressed and Decompressed Relative Weights

In Models 9 and 10, we added adjustments for serving low-income patients (using the log (1 + DSH) form) and for location in a rural area. The relative weights are compressed in Model 9 and decompressed in Model 10. Both models bring the PTC ratios for large urban and rural hospitals closer to 1.0. In Model 10, the PTC ratio for rural facilities is .99 compared to .87 in the base model. Despite making no adjustment for facilities located in large urban areas (which would be negative if adopted), the average PTC ratios for large urban facilities are only slightly higher (1.01) than for other urban facilities (.99) and rural facilities (.99). The average PTC ratios for units of acute care hospitals and freestanding hospitals would remain unchanged at 1.01 and .98, respectively. There is an increase in the PTC ratio for teaching hospitals that reflects the benefits for this group of making an adjustment for serving low-income patients. In particular, the average PTC ratio for the 24 hospitals with a resident-to-ADC ratio of between .10 and .20 increases from .99 in the base model to 1.02 in Models 9 and 10. The PTC ratio for the facilities with the highest teaching intensity increases slightly from .89 in the base model to .90 in Models 9 and 10.

With regard to the facility groupings based on percentage of low-income patients, the average PTC ratio for facilities with a low-income patient ratio below .30 is about 1.0. However, the average PTC ratio for the 19 facilities with a low-income percentage of .30 or higher is 1.07. These facilities would receive on average about 7 percent more than their costs using the log (1 + DSH) form of the low-income patient adjustment.

For the most part, there is little difference in PTC ratios for the various facility groups between Models 9 and 10. The major difference is in facility quartiles of CMI values. Overall, the average PTC ratios are closer to 1.0 using the decompressed relative weights in Model 10. The greatest difference can be seen in the lower quartile of CMI values, where the PTC ratio is 1.04 and 1.02 in Models 9 and 10, respectively. Facilities located in the Middle Atlantic region would have a slightly lower PTC ratio using the decompressed relative weights in Model 10 (1.010 versus 1.016 in Model 9), whereas those located in the Pacific region would have a higher average PTC ratio (1.038 versus 1.027 in Model 9).

The payment simulations support using the labor-related share to adjust for geographic differences in wage levels. Even though the fully specified regressions indicate the wage index is compressed, the average PTC ratio for the lowest wage index quartile is less than 1.0 (.998 in Model 10) and nearly equal to 1.0 for the highest quartile. It is slightly higher than 1.0 in the remaining inner quartiles.

Comparison of Low-Income Patient Adjustments

The remaining simulation models examine alternative formulations of the low-income patient adjustment. The impact of these alternatives is generally small across the facility groups, and the differences are concentrated in the low-income patient categories. We present more-detailed comparative information for these models relative to Model 10 in Table 7.18. In general, the patterns of profit and loss are similar across the models except in the group with a low-income patient ratio of .30 or higher.

Model 10A uses the LIP ratio instead of the DSH ratio. The model results in PTC ratios that are quite similar to those using Model 10. The patterns across winners and losers are also very similar. With the exception of one facility profiting under Model 10 and losing under model 10A, the same facilities would profit or incur a loss under either model. With one exception, the difference between the two models in average profit or loss per case in each low-income patient group is less than $25 per case. The exception is the nine facilities accounting for 38.4 percent of the cases in the group with an average low-income patient ratio of .30 or higher. These facilities would incur an average loss of $2,359 under Model 10 and $2,243 under model 10A.

Table 7.18

Payment Simulation Results by Low-Income Patient Ratio

	Low-Income Patient Ratio											
	Less than .10			.10 and less than .20			.20 and less than .30			.30 or greater		
Payment rates												
N facilities	416			213			46			19		
N cases	150,256			77,622			14,453			5,130		
Cost per case	$10,635			$12,255			$13,452			$14,430		
Payment ratio												
Model 10	1.0070			0.9826			0.9987			1.074		
Model 10A	1.0070			0.9828			1.000			1.077		
Model 10C	1.0059			0.9897			0.9913			1.028		
Model 10CT	1.0084			0.9864			0.9881			1.025		
Winners	No. of Facil-ities	% of Cases	Avg. Profit	No. of Facil-ities	% of Cases	Avg. Profit	No. of Facil-ities	% of Cases	Avg. profit	No. of Facil-ities	% of Cases	Avg. Profit
Model 10	208	59.4	$1,627	93	52.9	$1,590	21	53.1	$2,022	10	61.6	$3,204
Model 10A	209	54.6	$1,620	94	53.2	$1,586	21	53.1	$2,067	10	61.6	$3,193
Model 10C	206	53.8	$1,624	96	53.7	$1,656	21	53.1	$1,937	8	43.6	$3,496
Model 10CT	209	54.8	$1,629	95	53.6	$1,619	21	53.4	$1,890	8	43.6	$3,436
Losers												
Model 10	208	40.6	$-1,778	120	47.1	$-2,237	25	46.9	$-2,331	9	38.4	$-2,359
Model 10A	207	45.4	$-1,791	119	46.8	$-2,253	25	46.9	$-2,334	9	38.4	$-2,243
Model 10C	210	46.2	$-1,758	117	46.3	$-2,197	25	46.9	$-2,446	11	56.4	$-1,977
Model 10CT	207	45.2	$-1,775	118	46.4	$-2,227	25	46.9	$-2,484	11	56.4	$-2,019

The impact of using the log (.0001 + DSH ratio) form for the low-income patient adjustment (Model 10C) is concentrated in the group with a LIP ratio of .30 or higher. The average PTC ratio for these facilities is 1.03 (compared to 1.07 in Model 10) and reflects the flatter slope at the high ratios. Two facilities that would profit under Model 10 would incur a loss under Model 10C. Reflecting the additional two facilities in the loss category, the average loss per case is less under Model 10C ($1,977) than under Model 10 ($2,259). The change in the PTC ratio for facilities with a LIP ratio =>.10 and <.20 is modest but in the desirable direction. The ratio for these facilities is .9897 in Model 10C compared to .9826 in Model 10. The change in the ratio for the facilities with a LIP ratio =>.20 and <.30 is not in the desired direction. It is 1.000 in Model 10 and .9913 in Model 10C.

One of the disadvantages of the log (.0001 + DSH ratio) form of the low-income adjustment is the steep slope in the adjustment for the smallest patient ratios. It results both in significant payment differences for very small changes in the patient ratio and in a low national payment rate. For example, the national payment rate for the Model 10 simulation is $11,180. It drops to $8,072 in Model 10C. To counteract these disadvantages, we incorporate a 5 percent threshold into the log (.0001 + DSH ratio) form in Model 10CT.[19] The payment rate increases to $11,340. It is higher than in Model 10 because of the flatter slope of payment adjustment for the higher patient ratios. As expected, the few facilities with a ratio of less than .05 are advantaged and the PTC ratio for the <.10 group increases slightly. There are only small reductions in the PTC ratios across the remaining low-income groups.

---

[19] We fit the facility-level cost models using General Additive Models (GAM), which allow for more flexible relationships between the dependent variable and continuous independent variables. We found that the relationship between cost per case and serving low-income patients is not strictly linear and becomes positive around a low-income patient ratio of .02. However, the confidence interval at the lower ratios is quite large because of the small number of cases and provides little empirical support for using .02 as the threshold. We chose to use a higher threshold to avoid the steep increases at the lowest DSH ratios.

Supplemental Analyses

Type of Facility

A concern from the payment simulations is that freestanding hospitals have lower payment-to-cost ratios than do units of acute care hospitals. For example, the PTC ratio in Model 10 is .981 for freestanding hospitals compared to 1.014 for units. In Table 7.19, we explore whether there are differences in patterns of care that might account for the higher costs of freestanding hospitals. The average case weight in the small freestanding hospitals of 1.0937 compared to .9845 in the largest hospital size category is striking. It suggests that there are some small highly specialized freestanding hospitals with high case mix and costs per case. (The higher case mix in larger rehabilitation units relative to smaller rehabilitation units of acute care hospitals is consistent with the case mix pattern seen across general community hospitals.) These small freestanding hospitals have a lower proportion of tier 1 (more costly) comorbidities but a substantially higher percentage of tier 3 comorbidities relative to other freestanding hospitals. However, the rehabilitation units have a higher percentage of cases with comorbidities in each tier.

We do not know whether the higher incidence of comorbidities in rehabilitation units is attributable to treating sicker patients or coding practices. The higher average length of stay in freestanding hospitals is one indicator that coding practices may account for at least some of the differences:

- In the rehabilitation units, the average length of stay is 13.9 days, compared to an expected average length of stay of 15.7 days, or about 11 percent shorter.
- In contrast, the average length of stay for freestanding rehabilitation hospitals is 18.2 days, or about 16 percent longer than the expected average length of stay. The pattern of longer

Table 7.19

Comparison of Patterns of Care in Units of Acute Care Hospitals and Freestanding Hospitals

	All Facilities	Freestanding Hospitals				Units of Acute Care Hospitals			
	All	All	ADC <25	ADC 25-50	ADC >=50	All	ADC <10	ADC 10-25	ADC >=25
Average no. of cases	348	600	210	490	913	267	129	253	512
Average cost per case	11,330	11,547	14,649	11,733	11,091	11,172	11,668	10,964	11248
Standardized cost per case	11,696	11,940	14,036	12,433	11,404	11,520	12,512	11,504	11160
Case mix index	1.0014	0.9973	1.0937	1.0009	0.9845	1.0043	0.9902	0.9956	1.0208
% tier 1 comorbidities	0.66%	0.51%	0.37%	0.56%	0.49%	0.58%	0.58%	0.79%	0.81%
% tier 2 comorbidities	6.15%	4.95%	7.64%	4.59%	4.88%	7.01%	5.95%	6.72%	7.79%
% tier 3 comorbidities	11.54%	8.92%	12.52%	9.22%	8.35%	13.44%	13.51%	13.55%	13.26%
Average length of stay: all cases	15.7	18.2	17.4	17.7	18.5	13.9	12.6	13.4	15.1
Expected average length of stay	15.7	15.7	16.7	15.8	15.6	15.7	15.5	15.6	15.8
Transfer rate to hospital	0.078	0.087	0.095	0.086	0.086	0.071	0.061	0.067	0.080
Expected transfer rate	0.078	0.077	0.087	0.077	0.076	0.078	0.074	0.076	0.082
Transfer rate to SNFs	0.136	0.138	0.149	0.150	0.129	0.134	0.1432	0.136	0.128
Expected transfer rate	0.136	0.136	0.159	0.138	0.132	0.135	0.136	0.134	0.137
Short-stay transfers to hospital	0.062	0.065	0.077	0.064	0.065	0.060	0.054	0.058	0.066
Expected transfer rate	0.062	0.062	0.071	0.062	0.060	0.063	0.060	0.061	0.062
Short-stay SNF transfers	0.081	0.062	0.097	0.074	0.051	0.095	0.113	0.101	0.081
Expected transfer rate	0.081	0.082	0.098	0.083	0.079	0.081	0.082	0.080	0.082

lengths of stay than expected is consistent across the different
size categories of freestanding hospitals.

Despite the longer average lengths of stay, the freestanding
rehabilitation hospital rate for transfers to hospitals is higher than
expected. In contrast, the transfer rate from rehabilitation units to
acute care hospitals is lower than expected. Units may have better access
to services and physician consultants, thus allowing them to treat
patients with acute care needs without transferring them. On average,
freestanding hospitals also transfer a slightly higher percentage of
patients to SNFs than would be expected; however, the short-stay transfer
rate to SNFs is lower than expected. In contrast, the smaller
rehabilitation units have a higher-than-expected rate of short-stay
transfers to SNFs.

The TEP also questioned whether freestanding hospitals might have
higher overhead costs because they cannot allocate fixed costs to other
inpatient areas. In this regard, we note that freestanding hospitals have
higher capital-related costs per case than do units ($1,522 versus
$1,175), and capital constitutes a higher percentage of costs (12.1
percent versus 9.4 percent). As seen in Table 7.20, size is also a factor.
We will explore the issues related to overhead costs and facility
characteristics in greater depth as part of our refinement activities.

Table 7.20

Comparison of Capital-Related Costs, by Type and Size of Facility

Type and Size of Facility	No. of Facilities	ADC	Capital Cost	WI-and CMI-Adjusted Cost
All facilities	714	25	1,319	1,522
Units	545	18	1,175	1,275
ADC less than 10	168	7	1,498	1,458
ADC 10-24	274	16	1,106	1,269
ADC 25 or more	103	39	1,136	1,214
Freestanding hospitals	169	49	1,522	1,873
ADC less than 25	32	16	2,031	2,191
ADC 25-49	73	38	1,625	1,955
ADC 50 or more	64	77	1,405	1,791

Geographic Location

Differences in capital-related costs may also help explain the higher costs of rural facilities and the finding from our regressions that large urban hospitals are less costly than other facilities. We compare in Table 7.21 capital-related costs across geographic locations. Compared to the national average, large urban hospitals have lower capital-related costs per case and rural hospitals have higher costs per case. Capital-related costs are less that 10 percent of large urban hospital costs compared to 11 percent of costs for other hospitals. Again, size may also be a factor in explaining the higher costs of rural hospitals. This is an issue that we will examine as part of our refinement activities.

Table 7.21

Comparison of Capital-Related Costs, by Geographic Location

Geographic Location	No. of Facilities	ADC	Capital Cost per Case	WI-and CMI-adjusted Cost	% of Total Cost
All facilities	714	25	1,319	1,522	10.5
Large urban	329	28	1,308	1,368	9.8
Other urban	304	26	1,322	1,628	11.2
Rural	81	13	1,375	1,797	11.0

In Table 7.22, we explore whether there are differences in patterns of care that might account for the cost differences across geographic locations. The patterns of care are similar between large urban and other urban hospitals. Rural facilities, however, have a lower case mix index and a lower percentage of cases in each comorbidity tier. The average length of stay is similar across the three hospital groups:

- The average length of stay in rural facilities is slightly less than expected.
- The rural transfer rates to hospitals and SNFs are also lower than expected (although the short-stay SNF transfer rates are slightly higher than expected).

In short, there is no clear picture regarding how patterns of care might be affecting rural costs. The factor most likely to affect cost, longer length of stay, does not appear to be contributing to the higher costs

Table 7.22

Comparison of Patterns of Care, by Type of Location

	All Facilities	Large Urban	Other Urban	Rural
Average no. of cases	348	348	382	214
Average cost per case	11,330	11,651	10,947	11,888
Standardized cost per case	11,696	11,525	11,839	11,835
Case mix index	1.0014	0.9988	1.0046	0.9960
% tier 1 comorbidities	0.66%	0.66%	0.64%	0.59%
% tier 2 comorbidities	6.15%	6.75%	5.78%	4.65%
% tier 3 comorbidities	11.54%	11.58%	11.55%	11.21%
Average length of stay: all cases	15.7	15.6	15.8	15.4
Expected average length of stay	15.7	15.6	15.7	15.6
Transfer rate to hosp.	0.078	0.081	0.076	0.067
Expected transfer rate	0.078	0.078	0.077	0.074
Transfer rate to SNFs	0.136	0.135	0.137	0.132
Expected transfer rate	0.136	0.135	0.136	0.137
Short stay transfers to hosp.	0.062	0.065	0.061	0.055
Expected transfer rate	0.062	0.063	0.062	0.060
Short-stay SNF transfer rate	0.081	0.081	0.081	0.085
Expected transfer rate	0.081	0.081	0.082	0.082

of care in rural hospitals. Lower transfer rates may be a minor contributing factor.

The average length of stay for large urban facilities is as expected based on the mix of patients treated. Transfer rates to hospitals--both total and short stay--are slightly higher than expected and could be a possible factor in explaining their lower costs per case.

The TEP raised an issue regarding whether the facility adjustment for rural facilities should instead be an adjustment for serving rural patients. The underlying notion is that rural patients might cost more either because of the lack of available post-discharge placement opportunities or because they do not live near medical facilities and must reach a higher level of rehabilitation before discharge. (Unlike our current findings, our interim report showed that patients discharged from rural facilities had a slightly longer-than-expected length of stay.) It

was suggested that urban facilities that treat a substantial volume of rural patients should also benefit from a payment adjustment.

We sought to answer several questions: How do the costs of rural patients in urban hospitals compare to the costs of urban patients in those hospitals? And--vice versa--how do the costs of urban patients in rural hospitals compare to the costs of rural patients in those hospitals? We report our findings to this question in Table 7.23. After adjusting for case mix and the WI, we found the following:

- Rural patients cost more on average in urban hospitals than do urban patients ($12,544 versus $12,297).

- Urban patients cost more on average in rural hospitals than do rural patients ($15,094 versus $14,751).

Table 7.23

Comparison of Cost, by Hospital Location

Beneficiary Residence	Hospital Location	No. of Cases	Wage-Adjusted Cost	CMI-Adjusted Cost
Urban	Urban	372,177	$11,272	$12,297
Rural	Urban	68,956	$12,110	$12,544
Urban	Rural	3,068	$15,053	$15,094
Rural	Rural	29,192	$13,310	$14,751
All	All	473,393	$11,544	$12,502

Note: Based on combined 1998-1999 data. Beneficiary location was missing in approximately 3,500 cases.

In urban hospitals, rural patients cost about 2 percent more; in rural hospitals, urban patients cost about 2.3 percent more ($p < .0001$). However, the difference in cost by beneficiary residence is quite small compared to the difference in cost between urban and rural hospitals. For example, using Model 10, rural hospitals are 19 percent more costly than urban hospitals.

The second question we addressed was whether hospitals that cater to out-of-town patients tend to be more costly. To address this question, we matched 1998-1999 cases of rural beneficiaries in an urban hospital with an urban beneficiary in the same hospital and FRGC. We also matched urban beneficiaries in a rural hospital with a rural beneficiary counterpart in the same hospital and FRGC. We restricted the study to beneficiaries who were discharged to the community. We were able to match only 58 percent of

the rural cases. As seen in Table 7.24, matching reduces the difference to about $120 in urban hospitals, or just over 1 percent.

Table 7.24

Comparison of Cost, by Beneficiary Residence

Hospital Location	Total Out-of-Area Cases	Matched Cases	Rural Beneficiary Cost	Urban Beneficiary Cost
Urban	76,274	42,187	$11,056	$10,937
Rural	291,765	2,038	$13,515	$13,822

NOTE: Urban: $p = 0.0012$; rural: $p = 0.145$.

The third question we examined was what proportion of urban hospitals have a substantial proportion of rural patients. We report our findings in Table 7.25. Less than 50 percent of the patients in 96.2 percent of urban hospitals reside in rural areas. However, more than 50 percent of the patients in 4 percent of urban hospitals (39 hospitals) come from rural areas. These hospitals have a substantial rural patient load and may be at a slight disadvantage. However, if the rural patients in urban hospitals are similar to other rural patients at rural hospitals, the magnitude of the disadvantage is only 1 to 2 percent. Further, the cause may not be rural patients, but rather slightly higher needs in patients who travel a distance for care, which would also apply to urban patients in rural hospitals.

Table 7.25

Frequency of Proportion of Rural Patients in Urban Hospitals

Frequency of Rural Proportion	Cumulative Frequency	Cumulative Percentage
< .10	543	54.14
< .20	724	72.18
< .30	838	83.55
< .40	925	92.22
< .50	965	96.21
< .60	991	98.80
< .70	999	99.60
< .80	1,003	100.00

New Providers and Facilities with Undue Influence

Some of the results from our current analyses differ from our preliminary findings in the interim report. Two findings that are of particular concern involve the hospital wage index variable and the large urban variable:

- In our preliminary analyses for the interim report, we found that the wage index variable adjusted for the CMS Office of the Actuary's estimate for the labor-related share was not statistically different from the expected value of 1.0. In our current fully specified regression, the wage index is compressed. It is not statistically distinguishable from 1.0 in the payment regressions.

- Our preliminary analyses in the interim report found that the large urban variable was not significant. In the current regressions, the large urban variable is slightly significant using either the compressed or decompressed weights.

We used the added variable plot technique to investigate whether any hospitals might be unduly influencing the regression results of Model 2. Our purpose was to perform some preliminary analyses to identify the extent to which this might be an issue. A further analysis of the issue will be undertaken as part of the IRF PPS refinement activities. We report the results in Figure 7.2 for the hospital wage index. The line on the graph is the slope of the wage index coefficient. Each facility is represented by a circle, the size of which is determined by the facility's discharges. The facilities in the upper-right-hand quadrant are located in high-wage areas and have high costs. Six facilities in the upper-right-hand quadrant of the figure (all of which are located in New York) appear to influence the regression results. We found that when we controlled for these providers in the Model 4 regression, the wage index coefficient dropped from 1.183 (std. error = .105) to 1.053 (std. error = .110). As far as the New York facilities are concerned, we have no reason to believe their wage index values--which, at 1.448 for New York City and 1.378 for Suffolk-Nassau, are among the highest in the nation--are contributing to their "outlier" status. Their case-weighted average cost per case is
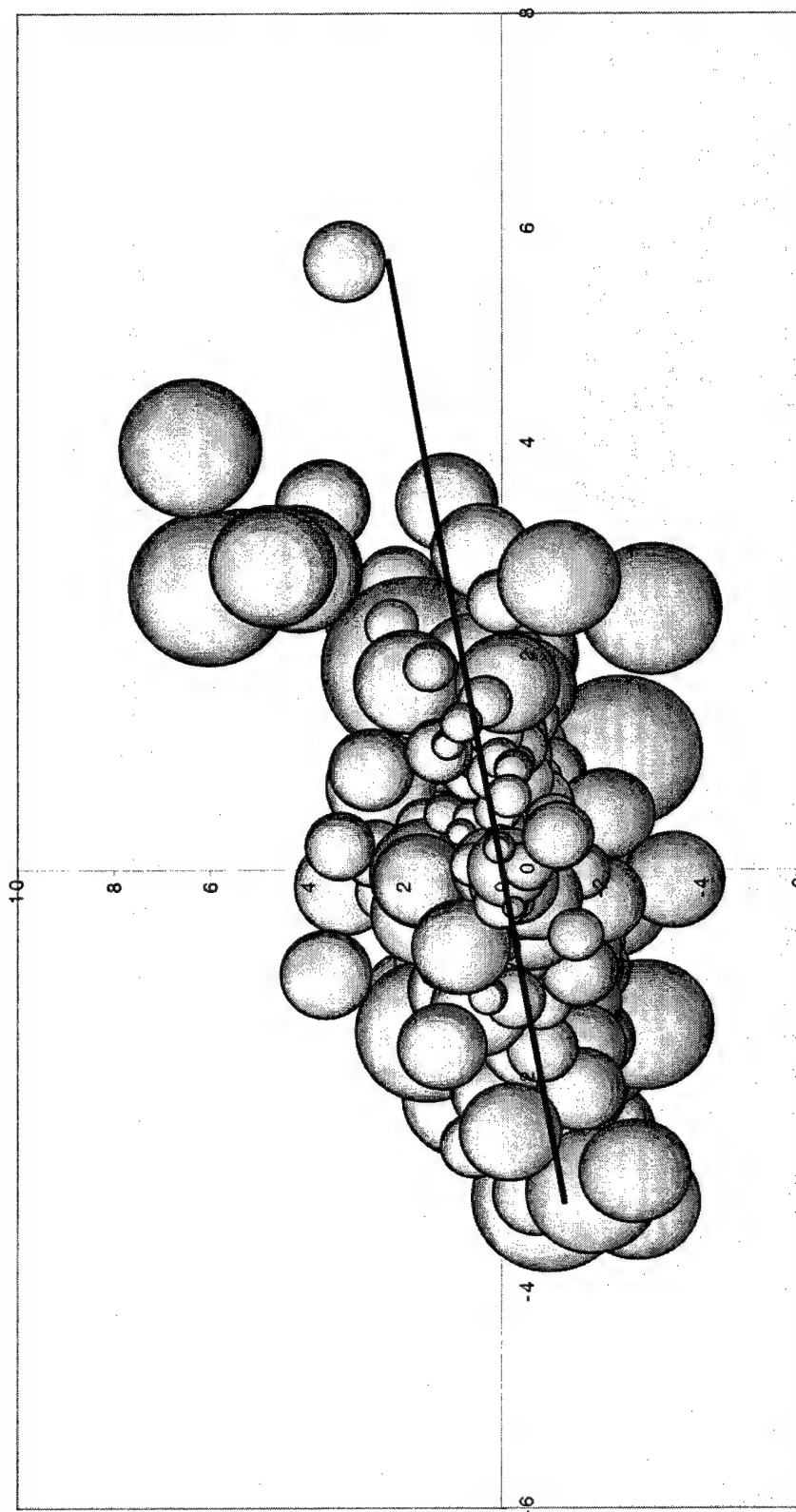
Figure 7.2--Facility Contributions to the Wage Index Coefficient (Model 2)

$16,300. After standardizing for payment parameters in Model 10, the adjusted cost for their 1999 cases is $11,773 compared to the national average cost of $10,844. The average for all New York hospitals in the analysis file with 1999 claims (31 facilities) is $10,372. Three of the six hospitals only recently began to use UDSmr and were not in our analysis file for the interim report. It is possible that inexperience with the FIM instrument may be contributing to their relatively low CMI value of .864.

Figure 7.3 shows that a group of low-cost large urban facilities might also be unduly influencing the large urban coefficient in the lower-right-hand quadrant (although there are also influential hospitals in the other quadrants).

The issue of influence of facility outliers on the regression coefficients and the characteristics of these facilities should be explored further during refinements to the IRF PPS facility adjustments. We determined, for example, that several outliers are new facilities that were still in their TEFRA base period. We subsequently identified 10 facilities in our database that were still in their TEFRA base period and did not have the same incentives as other facilities to control their costs. As Table 7.26 illustrates, these facilities look quite different from the rest of the analysis file and raise a policy issue regarding whether they should be included in the regression analysis. The facilities are larger and have both a higher cost per case and a higher case mix index. Eight of the 10 facilities are freestanding.

Table 7.26

Characteristics of New Facilities Compared to All Facilities

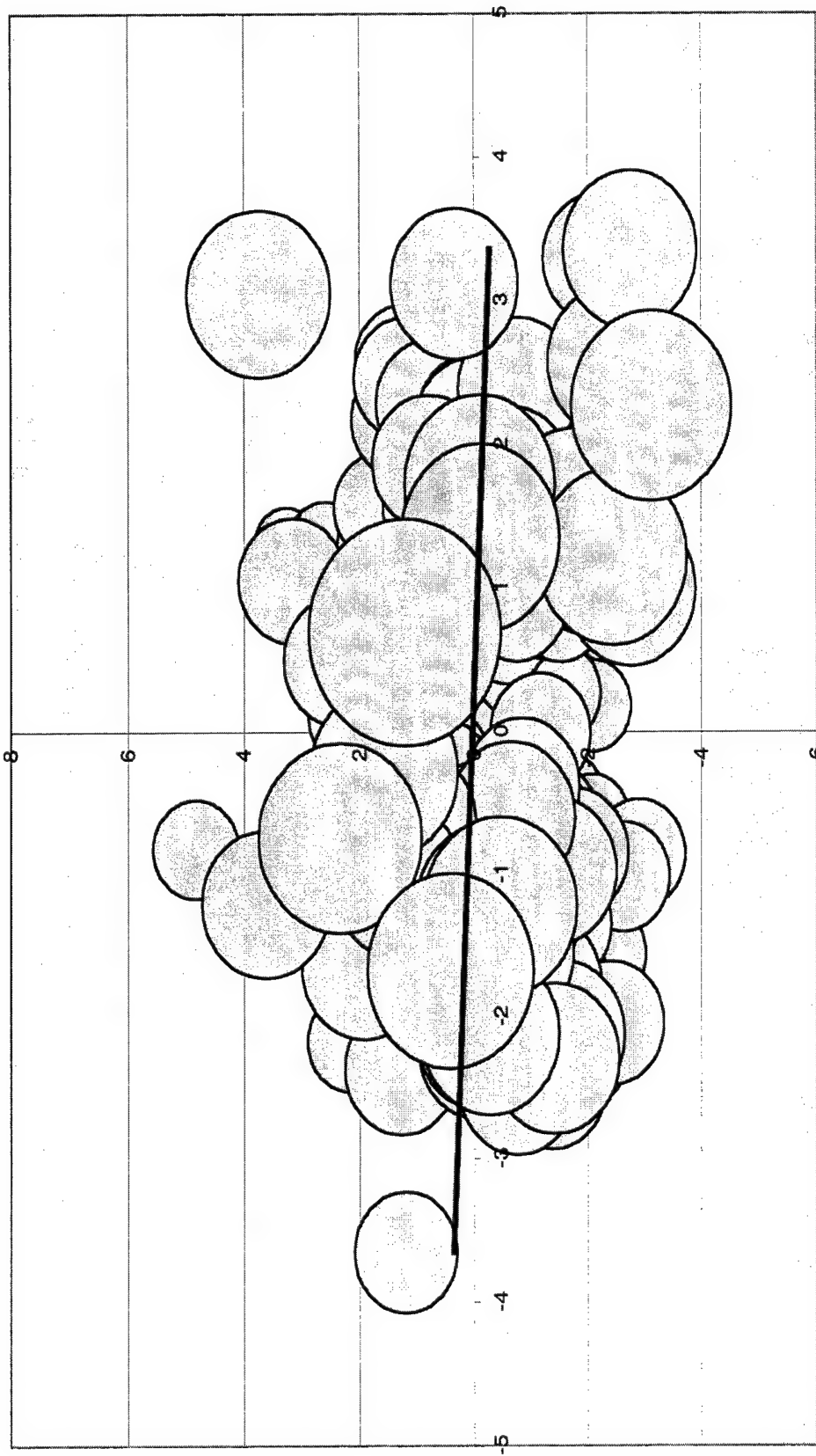|	No. of Facil- ities	Average Annual Cases	Cost per Case	CMI	% Free- standing	ADC
All facilities in file	714	320	12,073	0.999	24	25
New facilities	10	411	16,331	1.075	80	35

Figure 7.3--Facility Contributions to the Large Urban Coefficient (Model 2)

To understand the effect the new facilities might have on the facility adjustments, we controlled for them in a payment regression. Table 7.27 shows the effect of controlling for new facilities using the Model 10 regression specification. The dummy variable for the new providers is significant. There is only a slight effect on the rural coefficient. The coefficient for the DSH variable drops from .484 to .466. Using the coefficient as a payment adjustment would decrease the average adjustment from 5.7 percent to 5.1 percent. The regression should be considered illustrative only in that other outliers may be unduly influencing the coefficients for other variables. A systematic review of influential facilities and their impact on the payment adjustments should be undertaken once data are available for the full universe of rehabilitation facilities.

Table 7.27

Payment Regressions Showing the Effect of Controlling for New Facilities Using Cost per Case Standardized for WI and Decompressed CMIs with DSH and Rural as Other Explanatory Variables

	Model 10			
	Includes New Providers		Controls for New Providers	
Variable	Coef.	t-stat.	Coef.	t-stat.
Log (1 + DSH ratio)	0.484	4.146	.466	4.025
DSH gap-fill dummy	.0216	1.087	.026	1.303
Rural	0.175	5.303	.176	5.369
New provider dummy	---	---	.195	3.100
Intercept	9.317	567.447	9.312	71.636
R-squared	0.0601		.0875	

IMPLICATIONS FOR POLICY AND RECOMMENDATIONS

In this subsection, we describe our main empirical findings and their implications for specific aspects of the IRF PPS facility-level adjustments. If just the statutorily mandated wage adjustment were used as a facility-level adjustment, rural hospitals, hospitals with high proportions of low-income patients, and hospitals with high teaching intensity would have an average PTC ratio of less than 1.0. Below we discuss the facility-level adjustments that we believe are needed to improve payment equity.

Wage Index

When the wage index variable is adjusted for the CMS Office of the Actuary's estimate of the labor-related share, the coefficient is considerably higher than its expected value of 1.0 in the fully specified regressions. This is an issue that warrants further investigation as part of IRF PPS refinement activities. In the interim, we believe that using the CMS OACT estimate of the labor-related share is appropriate. Our preliminary investigation of potential causes for the wage index compression suggests that a small subset of facilities may be having an undue influence on determining the effect of wages on facility costs. The wage index coefficient decreases from 1.183 to 1.053 in Model 4 when the regression controls for these facilities. Moreover, the average PTC ratio for the hospitals in the highest wage index quartile is nearly 1.0, which suggests that using the labor-related share works well for these hospitals.

Rural Location

When cost per case is standardized for case mix and area wage differences, rural hospitals are 19 percent more costly than urban hospitals. We found that the rural adjustment based on the regression coefficient worked well. The PTC ratio across rural and urban hospitals using the regression coefficient as an adjustment is close to 1.0.

We found that rural patients in urban hospitals are about 1 percent more costly than their urban counterparts in the same hospital. About 4 percent of urban facilities draw more than half of their patients from rural areas. Although these hospitals might be at a slight disadvantage, the magnitude of the disadvantage is very small, and we do not believe an adjustment would be appropriate. Since urban patients who are in rural hospitals are also more costly than their rural counterparts, the underlying issue may be whether patients who travel out of area for inpatient care are more costly. Refinements in the patient classification system would be preferable to a facility adjustment for these small numbers of patients.

Large Urban Location

The payment regressions also indicate that large urban hospitals are about 4 percent less costly than other hospitals. This is another issue that warrants further investigation as part of IRF PPS refinement activities. In the interim, we do not recommend making an adjustment for location in large urban areas. In contrast to the rural coefficient, the large urban coefficient is only slightly significant. Moreover, our preliminary investigation of the large urban effect indicates a small subset of facilities may be unduly influencing the results. In addition, the higher costs of other urban hospitals relative to large urban hospitals may be attributable to coding practices and/or patterns of care rather than higher infrastructure costs. The other urban hospitals report a lower proportion of cases with comorbidities and have a slightly longer than expected average length of stay. We believe this issue should be reexamined after data from IRF PPS become available.

Type of Facility

The regressions indicate that freestanding hospitals are about 5 percent more costly than rehabilitation units of acute care hospitals. The PTC ratio for freestanding hospitals is .981. The issue of whether patterns of care contribute to the higher costs of freestanding hospitals is clouded by uncertainty over coding reliability. Based on their somewhat higher transfer rates, we could expect freestanding hospitals to have lower lengths of stay. Instead, they have a higher-than-expected length of stay. The issue should be examined after hospitals have responded to the payment incentives to improve coding practices and to deliver care more efficiently. In addition, the relationships between the type and size of facility and fixed costs should be examined during refinement activities.

Low-Income Patients

Our analysis of an adjustment for serving low-income patients focused on three issues: how commitment to serving low-income patient should be measured, whether the measure should be in a logged or non-logged form, and data quality.

Three potential measures of a facility's commitment to serving low-income patients could be implemented based on currently available data: the facility's low-income patient (LIP) ratio, its DSH ratio, or the

proportion of Medicare patients who are entitled to SSI. There is little
difference between the regression and simulation results using the LIP
ratio or the DSH ratio. We believe LIP is a better technical measure of a
hospital's commitment to serving low-income patients. However, there is
policy precedence for using the DSH ratio that should also be considered.
Either measure is preferable to one based solely on the rehabilitation
facility's SSI ratio.

The logged forms of the low-income patient measure result in higher
R-squareds than the non-logged form. The log (.0001 + ratio) form has the
highest R-squared (.065 compared to .060 for the log (1 + ratio) form).
However, the steep slope at the lowest patient ratios is a major drawback
to using this form to establish the payment adjustment. Establishing a
threshold to eliminate the effects of the steep slope would be arbitrary.
The form log (1 + ratio) is a reasonable alternative for avoiding the
steep slope without being arbitrary. Except for the highest low-income
patient category (.30 or higher), the payment simulations for both forms
are quite similar. Reflecting the steeper slope at the higher ratios
produced by the form log (1 + ratio), the PTC ratio for this group using
the form (1 + ratio) is 1.07 compared to 1.03 using the form log (.0001 +
ratio). However, the high average PTC ratio for the facilities serving the
highest proportions of low-income patients may not be a policy concern.

Our analysis of the data indicates that there may be considerable
underreporting of Medicaid days, which could affect the magnitude of the
low-income patient adjustment. The adjustment should be reevaluated with
data to be reported after the implementation of the IRF PPS. Even though
there are differences in Medicaid coverage across states, our analysis of
the UDSmr data suggests that broader measures of low-income patients do
not improve the explanatory power of the low-income variable. However,
this issue should also be reevaluated when data for the full universe of
rehabilitation facilities become available.

Teaching

The regressions show that teaching does not have a significant
effect on cost. We believe poor data quality may be contributing to the
regression results. In analyzing the cost report data, we found that the
salary allocations (and the resulting FTE resident counts that would

result from using those allocations) are often inconsistent with residents' assigned time. In addition, because patient census and bed capacity are relatively small in rehabilitation facilities, the measure of teaching intensity, particularly for units, can be sensitive to relatively minor changes in resident counts. When we account for these findings by using simply a measure of teaching/non-teaching as our teaching variable, teaching is still not significant. We recommended to CMS that more emphasis be placed on proper and consistent reporting of FTE counts in rehabilitation facilities and that the issue of whether there is a teaching effect on costs per case be reevaluated when better data become available.

Compressed/Decompressed Relative Weights

We also used facility-level regressions and payment simulations to evaluate the comparative effect of using the compressed or decompressed relative weights. We found that the R-squareds in the regressions using CMIs derived from the compressed relative weights are higher than those using CMIs derived from the decompressed relative weights. However, the payment simulations show that the hospitals in the lowest CMI quartile would have an average PTC ratio of 1.04 compared to an average PTC ratio of .9862 for hospitals in the highest CMI quartile using the compressed relative weights (Model 9). The ratios improve when the decompressed relative weights are used. The average ratio for the lowest CMI quartile falls to 1.02 while the average ratio for the highest CMI quartile increases to 1.0 (Model 10). There is no change in the average PTC ratio for hospitals in the inner CMI quartiles and little impact on other payment adjustments. We believe that the decompressed weights improve payment equity, and we recommended that they be used in the IRF PPS.

8. OUTLIERS

Outlier payments are additional payments, beyond the normal CMG payment, made for very expensive cases. Such payments can reduce hospitals' financial risk from a PPS (Keeler, Carter, and Trude, 1988) and should reduce the PPS incentive for hospitals to underserve very expensive cases. Also, by targeting payments to cases where the CMG payment is much lower than cost, they may help mitigate problems with the classification system. Because outlier cases are not paid at full cost, they cannot completely fix problems with the classification system; they can only provide some compensation. Outlier payments have several drawbacks: Because outlier payments are tied to costs, they may encourage hospitals to provide care that has less value than its cost. Further, because charges are used to measure costs, it might be possible for some hospitals to game their charging structure to obtain an unwarranted portion of the outlier payments. Finally, because behavior may change after implementation of the IRF PPS, there is uncertainty about the relationship between the outlier parameters and the total amount of funds that will be spent on outlier payments. Thus, the more outlier payments that are planned, the larger the possible difference between realized total payment and the budget neutrality target.

In this section, we update the analyses in our interim report concerning the amount of the outlier payments to be included in the IRF PPS. We use 1999 data to show that the basic conclusions of our interim report still hold with later data. In our interim report, we recommended that HCFA use a target for the total amount of outlier payments of 3 percent of total IRF PPS payments. Because the conclusions of our updated analyses are similar, we did not change our recommendation. CMS used a 3 percent target in both the NPRM and the final rule.

We simulated payments to evaluate the quantifiable outcomes of different targets for the amount of outlier payments: financial risk, the accuracy of the payment-to-cost match at the hospital level, and the payment-to-cost ratio for groups of hospitals. We then attempted to weigh the measured effects against other, unquantified effects in order to reach a recommendation.

The simulations also allowed us to summarize the performance of a PPS that would incorporate the elements of each of our major recommendations for policy:

- classification system
- payment for short-stay transfer cases
- HSRV weights after decompression
- the facility payment adjustment which uses the hospital wage index, rural location, and disproportionate share
- the level of outlier payment.

The classification system used for unusual cases is the same as the one described in Section 6. The HSRV weights from Table 6.6 were used. The facility adjustment is from Model 10 of Section 7. The simulations are based on 1999 data, and the method is discussed in Section 2.

OUTLIER PAYMENT FORMULAS

The simulated case payment includes a fixed-loss outlier policy with the loss threshold, L, chosen so that a fixed percentage of total IRF PPS payments would be made as outlier supplements. The outlier threshold for a case in hospital i in CMG k with comorbidity tier x is calculated as

$$T_{ik} = R * A_i * W(k,x) + L * A_i, \quad (8.1)$$

where, as before, R is the national conversion factor, A_i is the hospital's payment adjustment factor, $W(k,x)$ is the relative weight for the case,[1] and L is the policy parameter.

Because the hospital cost report is not available at the time of payment for the case, the cost of the case is estimated by multiplying the charges for the case by a cost-to-charge ratio (CCR). The CCR is specific to the hospital and is calculated from a recent cost report. For

---

[1] The relative weight for transfer cases is the weight for the CMG and tier times the min(1,(LOS + 0.5)/avg_los), where LOS is the case's length of stay and avg_los is the average LOS for the CMG and tier from Table 6.6. The fixed loss threshold is multiplied by the hospital's adjustment factor as is appropriate when costs follow the log normal distribution and the adjustment factor is based on measured cost, as we recommended in Section 7 and as CMS implemented.

convenience we used MEDPAR records in order to estimate the proportion of inpatient rehabilitation charges for each ancillary department.

The estimated cost of the case, C (= charges * CCR), is compared to the threshold. If C exceeds T_{ik}, then the case receives an outlier payment:

$$OUT = 0.8 * (C - T_{ik}).$$ (8.2)

This outlier payment is added to the CMG payment so that total payment for each case is

$$P = R * A_i * W(k,x) + OUT,$$ (8.3)

where OUT is 0 for non-outlier cases.

We allow outlier payments to occur for all kinds of cases-- including short-stay transfers, deaths, and atypical short-stay cases. Outlier payments are very rare for all these unusual cases except deaths.

SIMULATION PARAMETERS

We simulated policies with outlier payments equal to 0, 1, 2, 3, 4, and 5 percent of total IRF PPS payments. The first column in Table 8.1 shows the national conversion factor, R, for each of these runs. The conversion factor is much closer to the average payment per case than in the interim report because of the change in the form in which DSH enters the facility adjustment formula. The case-weighted average adjustment factor is 1.05 (rather than the 1.85 in the interim report and NPRM). As the percentage of money expended through outlier payments increases, the conversion factor declines to maintain budget neutrality. Thus each non-outlier case receives less payment as more money is spent on outliers. The outlier payment is a redistribution of payment from the non-outlier cases to the outliers.

Table 8.1 shows the loss threshold (L in Eq. (8.1)) and the loss threshold for the average case, which is L times the average value of the adjustment factor. For example, under the 5 percent outlier policy, a hospital with an adjustment factor equal to the average value would receive an outlier payment for a case that costs more than $7,571 above

its CMG payment amount. Under the 3 percent policy, losses at the same hospital in excess of $11,336 would receive outlier payments.

Table 8.1

Basic Statistics for Simulation Runs

Policy	Conver-sion Factor ($)	Average Payment per Case ($)	Loss Thresh-old ($)	Typical Adjusted Loss Thresh-old ($)	% of Cases with Some Outlier	Average Outlier Payment per Case ($)
No outliers	11,526	11,386	None			0
1% outliers	11,411	11,386	21,028	22,079	1.22	114
2% outliers	11,295	11,386	14,159	14,867	2.95	227
3% outliers	11,180	11,386	10,796	11,336	4.96	342
4% outliers	11,065	11,386	8,688	9,122	7.22	455
5% outliers	10,950	11,386	7,210	7,571	9.64	569

Note: Based on simulation of payment to 247,461 cases at 694 hospitals. Average cost per case was $11,386. Typical adjusted loss threshold is for a hospital with an average value of the hospital adjustment factor (1.05).

Average outlier payment per outlier declines and the number of outliers increases as the percentage of outlier payment increases. The percentage of outlier cases is greater than the percentage of payment to outliers. This suggests that the cost distribution is not very skewed. For example, in order to spend 5 percent of funds through outlier payments, we have to make outlier payments for 9.6 percent of all cases.

OUTLIER CASES

In all policies, outlier payments typically go to cases that would otherwise lose a substantial amount of money. Table 8.2 shows the costs and payments for cases receiving outlier payments under each policy. For example, the cases that receive an outlier supplement under the 1 percent outlier policy cost an average of $47,633 but received CMG payments of only $17,730. Although the outlier supplement added $9,377 to the total payment, the payment-to-cost ratio for these cases was still only 0.569.

Table 8.2

Cost and Payment for Outlier Cases in Each Outlier Policy

Outlier Cases	Number of Outliers	Average Cost	CMG Payment (Before Outlier)	Average Outlier Payment per Outlier	Average Total Payment	PTC Ratio
1% outliers	3,007	47,633	17,730	9,377	27,107	0.569
2% outliers	7,300	38,158	16,110	7,719	23,829	0.625
3% outliers	12,271	33,296	15,112	6,889	22,004	0.661
4% outliers	17,856	29,938	14,336	6,312	20,648	0.690
5% outliers	23,866	27,506	13,710	5,903	19,613	0.713

RISK

Table 8.3 shows the risk statistic--the standard deviation of annual profit expressed as a fraction of expected annual revenue. In the absence of outlier payments, a typical rehabilitation facility manager would face a situation where the difference between annual Medicare costs and revenue had a standard deviation of 3.32 percent of expected revenue due solely to a random draw of patients.

Table 8.3

Risk as a Function of Percentage of Outlier Payments

Policy And Hospital Group	All Hospitals		Freestanding Hospitals with ADC <25		Unit with ADC <10	
	Average Value of Risk Statistic	Improve- ment from Previous Policy (%)	Average Value of Risk Statistic	Improve- ment from Previous Policy (%)	Average Value of Risk Statistic	Improve- ment from Previous Policy (%)
No outliers	0.0332		0.0427		0.0496	
1% outliers	0.0307	7.5	0.0363	14.9	0.0453	8.7
2% outliers	0.0289	5.9	0.0325	10.4	0.0426	5.9
3% outliers	0.0274	5.2	0.0300	7.7	0.0405	5.1
4% outliers	0.0262	4.4	0.0281	6.4	0.0386	4.5
5% outliers	0.0251	4.2	0.0266	5.4	0.0370	4.2

The smallest hospitals are at the most financial risk from a PPS-- they will find it much harder to cover the large loss that could be caused by a very small number of very expensive patients. Consequently, Table 8.3 also shows the average risk faced by the smaller hospitals in our sample. There are 29 freestanding hospitals in our sample with an

average daily census of less than 25 patients. They constitute about 18 percent of the freestanding hospitals in our sample and care for about 2.6 percent of the patients in our sample. IRF units are typically much smaller than freestanding hospitals. A little more than a quarter of the units in our sample have an average daily census of 10 or less. These small units care for 8.6 percent of patients. Although the small units are on average much smaller than the small freestanding hospitals, they have only slightly greater risk because their patient populations tend to have more-homogeneous costs.

Financial risk declines as outlier payments increase, but at a decreasing rate of improvement. To help interpret the meaning of the risk statistic, we use the law of large numbers: Since annual loss is the sum of losses on each patient, annual loss has the normal distribution. Thus, we plot in Figure 8.1 the probability that a hospital will experience a loss in excess of 5 percent of revenues due to random causes. For a typical hospital, the probability would be about 7 percent if there were no outlier payments--so once every 15 years a manager could expect such a large problem. If instead the policy were a 3 percent outlier, the probability of such a loss drops to near 2 percent. The rate at which risk declines with increases in outlier payment decreases noticeably for typical hospitals and for small freestanding hospitals--but less so for small units.

ACCURACY AT THE HOSPITAL LEVEL

Table 8.4 summarizes the differences between average payments and costs at the hospital level. In the absence of outlier payments (and, as with all statistics here, in the absence of behavior change), the average case would go to a hospital where average costs differ from average payments by $2,055 per case. As outlier payments increase, this number decreases, showing that outlier payments are more likely to go to hospitals where average costs exceed payment. The root mean square error (RMSE) also shows the typical difference but gives more weight to hospitals that have large differences between their costs and payments. We can use the RMSE to calculate the percentage of variance in average cost across hospitals that is explained by payment, which is the Efron R-Squared. With no outlier payment, payment per case explains 38.6

percent of the hospital variation in cost per case. With a 3 percent
outlier payment, this percentage rises to 57.1 percent, and with a 5
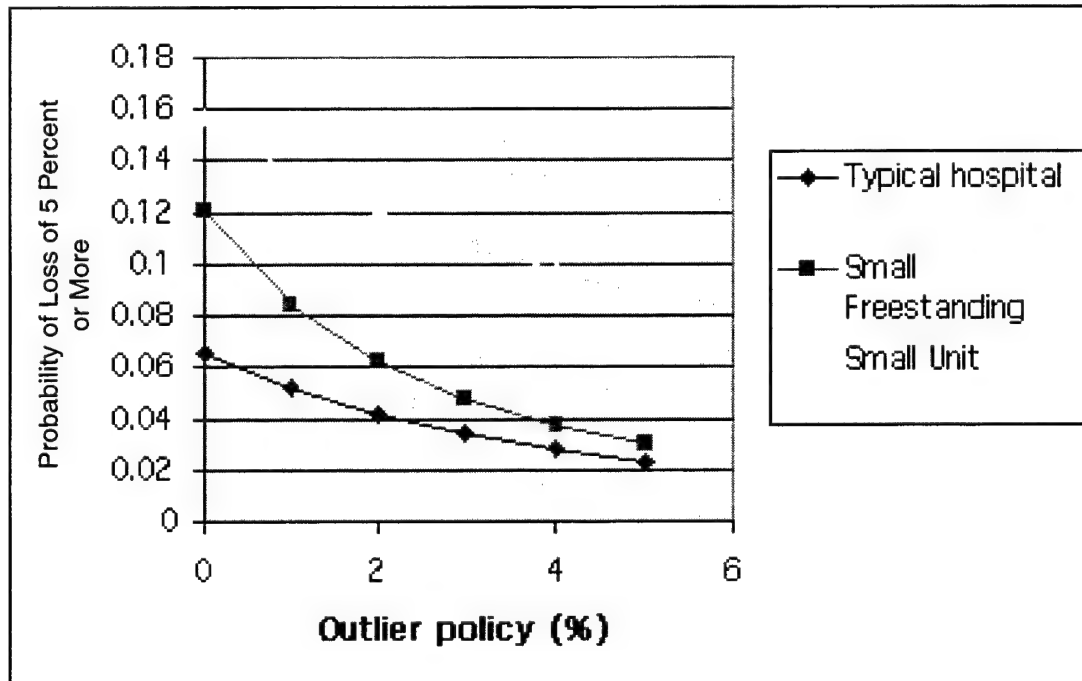percent payment, to 64.8 percent.

Figure 8.1--Probability of Loss of 5 Percent or More,
by Amount of Outlier Payment and Type of Hospital

Table 8.4

Accuracy of Hospital-Level Payments:
Root Mean Square Difference Between Payment and Cost and Mean Absolute
Difference Between Payment and Cost, by Policy

Policy	Mean Absolute Deviation	Improvement from Previous Policy (%)	RMSE	Efron R-Squared	Improvement from Previous Policy (%)
No outliers	2055		2861	0.386	
1% outliers	1973	4.0	2682	0.460	19.3
2% outliers	1889	4.3	2524	0.522	13.4
3% outliers	1810	4.2	2390	0.571	9.4
4% outliers	1736	4.1	2273	0.612	7.2
5% outliers	1665	4.1	2166	0.648	5.8

Note: The statistics are case-weighted.

Table 8.5 shows the payment-to-cost ratios for groups of hospitals for each outlier policy. For almost all groups of hospitals, the amount of outlier payment has little effect on the payment-to-cost (PTC) ratio for the group--typically less than 2 percent as the outlier payments go from 0 to 5 percent. This is because most outlier cases are widely distributed (as one would expect if they are really random events) and occur in all types of hospitals. This may be seen more directly in Table 8.6, which shows the amount of outlier payment per case.[2]

The last set of hospital groups in Tables 8.5 and 8.6 is defined based on CMI quartile. Although the pattern is not completely linear, hospitals with a CMI in the lowest quartile are slightly overpaid in these runs--by about 2 to 3 percent. However, the four quartiles are each closer to 1 than in our interim report or in the compressed-weight results. Outlier payments produce only a slight change in the PTC of any of these quartiles--less than one-half of 1 percent change in each of the three top quartiles as the policy changes from 0 percent outliers to 5 percent outliers. For the quartile of hospitals with the highest CMI, the PTC ratio is 1.001 with no outlier payment and 1.006 with a 5 percent outlier policy. At a 3 percent outlier policy, the PTC ratio for this group of hospitals is 1.005.

Using the final rule policy (3 percent payment), the largest deviation from 1 in the PTC ratio is for the small number of freestanding hospitals with an average daily census of less than 25. These hospitals have a PTC ratio of only 0.888. As discussed in Section 7, the teaching hospitals with a ratio of >=2 interns and residents per bed also have a low PTC ratio. The hospitals in all other groups have PTC ratios within 7.4 percent of 1; most within 3 percent.

---

[2] Rural hospitals and hospitals with a greater-than-average percentage of low-income patients ("DSH" facilities) are receiving close to their appropriate share of the outlier payments. Their PTC ratios vary little with an increase in outlier percentage. The PTC ratios for rural hospitals are close to 1 for all outlier policies; the ratios for hospitals with lower than average DSH (< 0.10) actually slightly decline toward 1 with increasing outlier payments, showing they are getting less than the policy percentage of outlier payments. Thus, those with higher DSH are getting more outliers.

Table 8.5

Payment-to-Cost Ratios for Hospital Classes, by Amount of Outlier Payment

Class	Number of Cases	Number of Hos-pitals	Cost per Case	Payment-to-Cost Ratio					
				FRG 3, 0% Outlier	FRG 3, 1% Outlier	FRG 3, 2% Outlier	FRG 3, 3% Outlier	FRG 3, 4% Outlier	FRG 3, 5% Outlier
All facilities	247,461	694	11,386	1.001	1.000	1.000	1.000	1.000	1.000
L urban	112,270	319	11,757	1.013	1.012	1.012	1.012	1.012	1.011
O urban	118,014	297	10,974	0.989	0.988	0.988	0.989	0.989	0.990
Rural	17,177	78	11,793	1.002	0.996	0.994	0.993	0.992	0.992
New England	12,015	26	13,058	0.943	0.951	0.956	0.961	0.964	0.967
Middle Atlantic	46,433	90	11,057	1.013	1.012	1.010	1.009	1.008	1.007
South Atlantic	48,420	106	10,667	1.065	1.060	1.056	1.052	1.050	1.047
East North Central	42,199	148	11,280	0.997	0.996	0.995	0.994	0.994	0.993
East South Central	21,029	37	11,077	1.031	1.026	1.023	1.022	1.020	1.019
West North Central	14,896	60	10,981	0.970	0.969	0.970	0.971	0.973	0.974
West South Central	39,885	118	11,707	0.916	0.918	0.923	0.928	0.932	0.936
Mountain	9,988	38	10,262	1.033	1.032	1.034	1.034	1.035	1.035
Pacific	12,596	71	14,994	1.037	1.037	1.038	1.038	1.037	1.037
Unit of acute-care hospital, size	142,504	529	11,265	1.017	1.015	1.015	1.014	1.013	1.013
Less than 10	21,281	164	11,819	0.937	0.940	0.944	0.947	0.950	0.952
10-24	67,387	262	10,955	1.025	1.021	1.018	1.017	1.015	1.014
25 or more	53,836	103	11,433	1.040	1.039	1.039	1.038	1.037	1.036
Freestanding hospital size	104,957	165	11,551	0.980	0.980	0.981	0.981	0.982	0.983
Less than 25	6,541	29	14,711	0.861	0.870	0.880	0.888	0.896	0.903
25-49	38,265	73	11,738	0.942	0.943	0.946	0.949	0.951	0.954
50 or more	60,151	63	11,088	1.024	1.021	1.019	1.017	1.016	1.015
Low-income percentage									
Less than 10%	150,256	416	10,635	1.012	1.009	1.008	1.007	1.006	1.005
10-19%	77,622	213	12,255	0.979	0.980	0.981	0.983	0.984	0.985
20-29%	14,453	46	13,452	0.992	0.994	0.996	0.999	1.000	1.002
30% or more	5,130	19	14,430	1.085	1.079	1.076	1.074	1.072	1.071
Ratio of teaching staff patients									
0	204,191	593	11,156	1.004	1.003	1.002	1.003	1.003	1.003
0-.09	29,679	61	12,166	0.984	0.985	0.986	0.986	0.986	0.986
.10-.19	10,672	24	12,971	1.034	1.029	1.026	1.024	1.021	1.020
.20 or more	2,919	16	13,722	0.880	0.886	0.893	0.898	0.902	0.906
Alaska/Hawaii	781	2	10,383	1.277	1.265	1.253	1.242	1.231	1.221
cmi q1	71,020	177	9,482	1.033	1.029	1.025	1.022	1.019	1.017
cmi q2	69,477	171	11,179	0.986	0.986	0.987	0.988	0.989	0.989
cmi q3	67,239	175	12,018	0.989	0.988	0.989	0.990	0.991	0.992
cmi q4	39,725	171	14,083	1.001	1.002	1.004	1.005	1.005	1.006

Table 8.6

Outlier Payment Amounts for Hospital Classes, by Amount of Outlier Payment

Class	Number of Cases	Number of Hospitals	Cost per Case	Outlier Payment Amounts ($)					
				FRG 3, 0% Outlier	FRG 3, 1% Outlier	FRG 3, 2% Outlier	FRG 3, 3% Outlier	FRG 3, 4% Outlier	FRG 3, 5% Outlier
All facilities	247,461	694	11,386	NA	114	228	342	455	569
L urban	112,270	319	11,757		124	243	358	472	586
O urban	118,014	297	10,974		112	225	339	453	567
Rural	17,177	78	11,793		58	148	252	362	477
New England	12,015	26	13,058		258	451	628	799	961
Middle Atlantic	46,433	90	11,057		115	213	311	409	508
South Atlantic	48,420	106	10,667		64	136	214	296	382
East North Central	42,199	148	11,280		109	214	319	425	531
East South Central	21,,029	37	11,077		68	154	248	346	448
West North Central	14,896	60	10,981		105	224	345	467	587
West South Central	39,885	118	11,707		143	304	465	622	775
Mountain	9,988	38	10,262		111	230	344	456	566
Pacific	12,596	71	14,994		180	347	504	654	803
Unit of acute hospital, size	142,504	529	11,265		106	215	323	430	537
Less than 1	21,281	164	11,819		167	322	471	614	755
10-24	67,387	262	10,955		75	163	254	349	446
25 or more	53,836	103	11,433		121	238	350	458	565
Freestanding hospital, size	104,957	165	11,551		125	245	367	491	613
<25	6,541	29	14,711		273	545	797	1036	1263
25-49	38,265	73	11,738		146	288	431	572	711
=>50	60,151	63	11,088		95	184	280	379	480
Low-income percentage									
< 10%	150,256	416	10,635		89	185	282	380	480
10-19%	77,622	213	12,255		155	293	429	563	697
20-29%	14,453	46	13,452		166	337	501	656	807
=>30%	5,130	19	14,430		83	196	321	452	591

Table 8.6 (cont.)

Class	Number of Cases	Number of Hospitals	Cost per Case	FRG 3, 0% Outlier	FRG 3, 1% Outlier	FRG 3, 2% Outlier	FRG 3, 3% Outlier	FRG 3, 4% Outlier	FRG 3, 5% Outlier
				Outlier Payment Amounts ($)					
Teaching: interns and residents to ADC									
0	204,191	593	11,156		108	218	330	443	556
> 0.05	29,679	61	12,166		153	286	410	529	646
0.05-0.15	10,672	24	12,971		86	182	283	388	498
=< 0.15	2,919	16	13,722		235	453	641	817	986
Alaska/ Hawaii	781	2	10,383		0	13	27	49	74
cmi q1	71,020	177	9,482		61	125	193	264	339
cmi q2	69,477	171	11,179		119	242	363	482	600
cmi q3	67,239	175	12,018		119	248	379	511	641
cmi q4	39,725	171	14,083		191	352	506	657	806

9. CONVERSION FACTOR

This section describes the determination of a budget-neutral conversion factor. The *budget-neutral conversion factor* is a national payment factor that, along with facility adjustments, case weights, and outlier adjustments, determines the IRF PPS payment for each case.

Congress mandated that the payment system in FY 2002 must be designed to be *budget neutral*--that is, estimated IRF PPS payments per case must not exceed estimated TEFRA payments per case. The IRF PPS will be phased in over two years. During cost reporting periods that begin after January 1 in federal FY 2002, a hospital's payment will be a blend of one-third the payment under TEFRA and two-thirds the payment under the IRF PPS. In the Benefits Improvement and Protection Act of 2000, Congress allowed hospitals to choose to be fully covered by the IRF PPS in FY 2002 rather than by the blend. The ability to make this choice was not to affect the conversion factor, however--i.e., the budget neutrality calculation was to assume all hospitals would be paid from the blend. For cost reporting periods that begin in FY 2003, the payment will be entirely through the IRF PPS.

The budget-neutral conversion factor computed in this section assumes that the cases at each hospital in FY 2002 will be similar to the hospital's cases in 1999 and that the OACT projections of TEFRA payments in FY 2002 are correct. The computation accounts for facility adjustments and outlier payments. CMS's actuary adjusted this conversion factor by decreasing it 1.16 percent in order to account for a decline in TEFRA payments due to changes in expected practice patterns based on experience from other PPSs. Although we believe that the comorbidities and FIM scores in our data contain errors, no adjustment was made to the conversion factor to account for improved coding.

The section begins with the rationale for our analysis strategy. We then discuss data used for computation of the conversion factor. Finally we evaluate several methods for handling cases at hospitals for which FIM data (and thus hospital CMI) are missing or incomplete and report the resulting conversion factors.

MOTIVATION

Since the conversion factor multiplies all IRF PPS case payments, it is key to meeting the congressional budget neutrality mandate. This section assesses several alternative methods for determining the national conversion factor under the IRF PPS.

Under the IRF PPS, each discharge is classified into a case mix group (CMG), thereby assigning a weight (CMG weight) to the case that is proportional to the cost of a typical case in the CMG. The case mix index (CMI) for a rehabilitation facility is computed as the average of the CMG weights for all cases at the facility. The CMI reflects the expected resource use for patients at the facility and can be regarded as a measure of the relative cost of each hospital's cases.

Congress mandated that, in FY 2002, estimated IRF PPS payment per case must equal estimated payment per case under the current (TEFRA) payment system. CMS's Office of the Actuary used trends to predict the average TEFRA per-case payment for each hospital and the number of cases at each hospital. We therefore know the target national average payment per case. Since outlier payments are planned to be a specified fraction (3 percent) of total payment, we also know the target for average non-outlier IRF PPS payments.[1]

Case payment under the new IRF PPS is a function of the national conversion factor, a known facility-specific adjustment factor, and CMG weight. Thus, the average payment per case minus outlier payments under the new system depends on the CMI at each hospital. If we know the CMI at each hospital, we can determine the conversion factor so as to meet the congressional mandate.

CMI is missing when FIM data are missing, and we are missing FIM data for many rehabilitation facilities. Further, the FIM data that we have were collected voluntarily, and we are also missing FIM data for some discharges at hospitals that did provide data.

In the NPRM, HCFA proposed a simple model in which the conversion factor is calculated for a sample of hospitals for which FIM data and other necessary data are available. This conversion factor would then be

---

[1] Attaining budget neutrality is, of course, also dependent on attaining outlier payments close to the target amount.

applied to the universe of hospitals. This method requires the assumption that the sample and the universe are similar in their relationship between TEFRA payment and CMI and facility adjustment factor. Our sample of hospitals is not entirely representative of the universe on several dimensions; for example, our sample contains a greater percentage of freestanding hospitals than does the universe. Because of the dissimilarity of the universe and sample with respect to characteristics of interest, we may be able to strengthen our ability to predict CMI by incorporating characteristics into our model that might be predictive of CMI. We would like to utilize the relationships between CMI and case-level and hospital-level characteristics to predict CMI for hospitals and cases with missing FIM data.

The immediate objective of the analyses reported here was to understand whether, in terms of the payment factor, we could do better than the simplest model in predicting CMI by incorporating case- and hospital-level characteristics into models of CMI. We examine and describe several CMI prediction models. We evaluate each of them on the basis of prediction error for estimating the conversion factor under the new payment system given the budget neutrality requirement.

Finally, we examine the national conversion factor implied by each model. In doing so, we assume no behavioral response to the IRF PPS in our estimates. Rather, we attempt to estimate the conversion factor under the assumption that hospitals will care for similar cases under the IRF PPS in the same way they cared for them in CY 1999 and will report them in a similar way.

DATA

The primary data sources for this analysis are the MEDPAR records and our matched FIM-MEDPAR database. We used only rehabilitation cases discharged in CY 1999. We further restricted our analysis to discharges from hospitals that appear in the OACT file that contains estimates of TEFRA costs and payments in FY 2002.

The outcome variable of interest, hospital CMI, was computed from the observed CMG weight for each case in the matched MEDPAR-FIM database. For most cases the weight was merely the weight assigned to the CMG. However, if a case was a transfer and its LOS plus one half-day

was less than the expected LOS for the patient's CMG category, the CMG weight was multiplied by the observed LOS plus the half-day per diem divided by the average LOS for that CMG. This modification fixes the weight of a transfer case to be no more than the case weight for its CMG. The weights and expected LOS used here were given in Table 6.6. They reflect the definition of CMGs developed in Sections 3, 4, and 5 and summarized in Section 6, and the decompression algorithm.

We would like to predict case weight for cases with missing FIM data so that these cases and hospitals can be included in our calculation of a conversion factor. We predicted CMI at hospitals based on hospital characteristics and on MEDPAR records of rehabilitation discharges and acute care stays that immediately preceded the rehabilitation stay. Hospital-level characteristics examined include

- an indicator of freestanding/exempt unit status
- a rural/urban indicator
- an indicator of a large (i.e., greater than 1) wage index.

Independent case-level variables examined include

- an indicator of whether the case was transferred to a skilled nursing facility (SNF)
- indicators of length of stay (LOS)
- an indicator of transfer status (including transfer to SNF or hospital)
- LOS and transfer interactions.

The transfer indicators came from the discharge destination field on the MEDPAR records. Of course, the case weights use transfer status from the FIM data, which we believe to be more accurate. Predicted RIC and predicted comorbidity were also entered into the model as independent variables; the method used for predicting RIC and comorbidity is described in the following subsection.

To assess how well we can predict CMI using alternative types of data and models, we then examined the accuracy of the conversion factor derived from the CMI predictions.

PREDICTING CASE MIX WITHOUT FIM DATA

Information about the case weight of a rehabilitation stay can be gleaned from the MEDPAR record for the rehabilitation stay and from the

MEDPAR record for the acute stay that preceded the rehabilitation stay.
Acute stays were included only when the discharge date was in the month
prior to the rehabilitation admission. (In 95 percent of cases, the
acute stay discharge date was the rehabilitation stay admission date.)

We created an estimate of the rehabilitation impairment category
for each rehabilitation stay that does not depend on the FIM data. It
was produced by combining information about the surgical procedures
performed during the acute stay, the principal diagnosis of the acute
stay, and all the diagnoses during the rehabilitation stay. We created
three variables, one for each source of information, and then developed
a rule to combine them.

The surgical procedures performed during the acute stay were used
to assign variable PREDRIC1. Heart transplants, valve procedures, and
CABG were assigned to PREDRIC1 = 14 (cardiac); liver transplants and
kidney transplants, to PREDRIC1 = 20. If the patient received a hip
replacement or a knee replacement, PREDRIC1 was set to 8 (joint
replacement of the LE). If the patient had a LE amputation (except for
the toe or foot) (codes 84.10 84.13 84.14 84.15 84.16 84.17 84.18 84.19)
or a "stump revision" procedure (84.3), PREDRIC1 was set to 10.
Otherwise, if the patient had an amputation of the toe, foot, or upper
limb except finger or thumb (codes 84.00 84.03 84.04 84.05 84.06 84.07
84.08 84.09 84.11 84.12), PREDRIC1 was set to 11. For selected back
procedures, including spinal fusion, PREDRIC1 was set to 5 (non-
traumatic spinal cord injury, or NTSCI).

The diagnoses on the rehabilitation stay need not include the
primary impairment and do not distinguish between the primary impairment
and comorbidities. Nevertheless, the diagnoses are useful for
identifying cases in the amputation RICs (any diagnoses in "V4974,"
"V4975," "V4976," "V4977," "V4978," or "V4979" were assigned to PREDRIC2
10; if none of these diagnoses appeared, any diagnosis in V496X or
"V4971","V4972", "V4973" was assigned to RIC 11) and cases with
Guillain-Barré (diagnosis 357.0), paraplegia, and major multiple trauma.
Only cases with third-degree burns were assigned to PREDRIC2 = 21
(burns). Paraplegic cases were divided between TSCI (PREDRIC2 = 4) and
NTSCI (PREDRIC2 = 5), depending on whether the primary diagnosis for the
preceding acute stay indicated trauma. Variable PREDRIC2 was assigned to

RIC 17 or RIC 18 based on the presence of multiple trauma codes and on whether injury to brain or spinal cord could be assigned.

We used a UDSmr list as a starting place for defining a mapping from the principal diagnosis for the preceding acute stay to the impairment code. This UDSmr list contained diagnoses that are appropriate to each impairment code. Several diagnoses were included under multiple impairment codes, and we selected one impairment code per diagnosis at most. We made several other changes at the code level. For example, many of the ICD-9-CM codes listed under the amputation impairment group codes were conditions that might cause an amputation but often do not. These were removed from the list. The final mapping from principal diagnosis to RIC was used to create variable PREDRIC3.

Finally, a single RIC was assigned to the case using the following algorithm:

- If PREDRIC3 (from the principal diagnosis of the acute stay) is RIC 7 (LE fracture), then PREDRIC = 7.
- Otherwise, if PREDRIC1 (from surgical procedure of the acute stay) is known, then PREDRIC = PREDRIC1.
- If PREDRIC is not yet assigned and PREDRIC2 (from rehab diagnosis) is 10, 11, or 21, then PREDRIC = PREDRIC2.
- If PREDRIC is not yet assigned and PREDRIC3 is known, PREDRIC = PREDRIC3.
- If PREDRIC is not yet assigned and PREDRIC2 (from rehab) is known (i.e. 4, 5, 6, 17, 18, or 19), PREDRIC = PREDRIC2.

Finally, if the case has LOS <= 3 and the MEDPAR rehabilitation record does not indicate a transfer, the case is predicted to be in the atypical short-stay CMG. If the MEDPAR rehabilitation record indicates an in-hospital death, the case is predicted to be (and actually is) in one of the death CMGs.

Cases that are not assigned a predicted RIC from this algorithm are further distinguished according to whether or not we found a record of a preceding acute care stay.

Table 9.1 shows the number of cases in each RIC and the percentage of cases that are predicted correctly in each RIC. It also breaks the remaining cases into those for which an incorrect prediction is made, those for which no hospitalization is found, and those for which a

Table 9.1

Accuracy of RIC Prediction and Average Cost and Case Weight, by RIC
(1999 data)

| RIC | Discharges | Percentage of Cases | | | | Average Cost ($) | | Average Case Weight |
		Predicted Correct	Predicted Incorrect	No Hospi-tal	Other Not Pre-dicted	In RIC	In Predicted RIC	In Predicted RIC
1	55,497	70.71	13.63	6.41	9.24	15,030.50	15,203.81	1.38
2	3,235	52.15	30.45	6.92	10.48	14,718.82	14,438.62	1.33
3	5,533	33.22	44.23	4.72	17.84	14,261.75	13,232.71	1.22
4	1,472	9.51	62.02	18.34	10.12	18,260.10	21,205.49	1.67
5	7,763	63.07	16.91	6.99	13.02	12,443.42	11,204.10	1.03
6	12,156	32.06	43.02	8.26	16.66	12,249.75	12,631.14	1.10
7	29,251	84.62	8.64	3.30	3.45	10,991.41	11,060.39	0.99
8	49,693	86.11	10.90	1.47	1.52	7,679.23	7,513.52	0.68
9	12,424	30.84	40.55	11.92	16.69	9,967.22	10,541.07	0.97
10	8,515	86.87	5.85	4.31	2.97	13,069.92	13,102.08	1.19
11	935	54.65	37.86	2.67	4.81	11,708.48	12,228.50	1.10
12	6,275	2.14	66.41	16.24	15.22	10,213.18	8,983.95	0.86
13	3,048	6.00	52.49	22.15	19.36	10,459.08	11,337.30	0.98
14	10,677	66.31	21.13	1.82	10.74	9,882.55	11,100.06	1.01
15	7,210	28.32	47.67	5.71	18.29	11,735.64	11,378.95	0.96
16	3,738	2.27	57.84	18.49	21.40	9,369.66	10,463.24	0.94
17	2,360	12.03	67.84	7.42	12.71	12,266.08	12,127.36	1.06
18	387	10.34	74.94	5.94	8.79	16,697.23	17,277.91	1.35
19	418	72.73	20.33	2.15	4.78	17,928.02	17,211.37	1.45
20	29,650	42.26	24.46	7.48	25.80	11,091.63	11,759.67	1.07
21	165	66.67	20.61	6.06	6.67	16,551.79	16,141.19	1.37
50	5,215	95.69	3.55	0.15	0.61	2,117.91	2,092.73	0.36
51	1,191	100.00	0.00	0.00	0.00	11,723.27	11,723.27	0.92
Total	256,808	62	22	6	10	11,308.30	11,164.22	1.01

hospitalization was found but the principal diagnosis has not been
mapped into any RIC.

Our predictions are quite accurate for three of the largest RICs:
stroke (RIC 1), fracture of the LE (RIC 7), and joint replacement of the
LE (RIC 8). We also do very well in RICs 10 (amputation of the LE) and
19 (Guillain-Barré). We predict correctly in over half the cases in 12
of the 23 bins. Overall we correctly predict 62 percent of the cases.

It is likely that each of the elements in the algorithm could be
improved somewhat. However, some inherent limitations in the data cannot
be overcome. The MEDPAR data are likely to be in error for some cases.
ICD-9-CM codes sometimes do not adequately distinguish the real reason
for rehabilitation. Further, some errors are inevitable in the FIM
coding of primary impairment in our database. Consider cases that enter

rehabilitation following major joint replacement with a principal diagnosis for the acute stay of hip fracture. According to UDSmr, the choice between RIC 7 and RIC 8 for these patients depends on the nature of the surgery. Emergency treatment is required following hip fracture; this emergency puts the patient at higher risk and therefore requires an assignment to RIC 7.[2] Nevertheless, a substantial number of these cases are assigned to RIC 8 (31 percent in the 1996-1997 data); these probably reflect errors in coding the principal diagnosis in the acute stay (many alternative codes would result in the same DRG assignment) as well as errors in the impairment code. As another example, we have a few thousand cases that were assigned to RIC 12 (osteoarthritis) when they entered rehabilitation following replacement of a major LE joint for a principal diagnosis of osteoarthritis. Although the rehabilitation might not be related to the surgery in some cases, we expect that the vast majority of these cases actually belong in RIC 8 and thus the high error rate in RIC 12 reflects in part a substantial amount of miscoding on the FIM.

The last two columns show that costs vary substantially across both actual and predicted RIC. Stroke cases cost about twice as much as RIC 8 cases (joint replacement of the LE), and several of the smaller RICs cost even more. Thus, insofar as hospital case mix varies across RICs, the CMI might be well predicted from a prediction of RIC.

We also predicted the comorbidity tier for each case. If PREDRIC had been assigned, we used the comorbidity list from the rehab stay with exclusions--as if RIC = PREDRIC--to assign the predicted tier. If PREDRIC was not assigned, the predicted tier was assigned a missing value; we had 37,655 cases in the sample (16.0 percent of cases) for which predicted tier was missing. As may be seen in Table 9.2, we can predict comorbidity tier quite well.

---

[2] A revision of the artificial joint may be elective and therefore appropriately assigned to RIC 8. However, only 1 percent of the hip fracture cases assigned to RIC 8 following major joint replacement recorded a hip revision on their acute stay record (procedure code 8153).

Table 9.2

Accuracy of Comorbidity Prediction by Comorbidity Value for Cases with Predicted RIC (1999 data)

Tier	Discharges	Percentage of Cases				Average Cost ($)		Average Case Weight
		Pre-dicted Correct	Pre-dicted Incorrect	No Hos-pital	Other Not Predicted	In Tier	In Predicted Tier	In Predicted Tier
1	1,613	73	3	5	19	18,432.66	18,722.43	1.41
2	16,088	69	11	4	15	13,700.36	13,978.90	1.17
3	29,811	69	13	5	14	12,683.08	12,632.99	1.05
None	209,296	81	4	6	9	10,508.19	10,411.67	0.89
Total	256,808	78	5	6	10	11,010.41	10,861.78	0.93

From the rehabilitation MEDPAR record, we can learn whether the patient died in the hospital, whether the patient was discharged to a hospital or SNF/nursing home, and the patient's LOS. We also used this information in predicting case mix, as detailed below.

MODELS FOR PREDICTING CASE MIX

To understand whether it is worthwhile to use the MEDPAR records to predict CMI for a hospital and, if so, how best to predict CMI, we examined four models:

Model 1. Use the case-weighted mean of the hospital CMI from our sample as the predicted CMI for the remaining hospitals in the universe.

Model 2. Predict CMI by stratifying the sample with respect to hospital characteristics.

The motivation for this analysis is that we know our sample and universe differ on the basis of hospital characteristics, such as the percentages of freestanding rehabilitation hospitals and exempt units. CMS will pay for certain hospital-level characteristics, such as wage index and rural status, and all of these factors were significantly related to CMI in exploratory linear regression analyses. We stratified the hospitals in our fitting sample on the basis of whether a rehabilitation facility was freestanding or a unit of a hospital, whether the facility was located in a rural area, and whether the rehabilitation adjusted wage index was greater than 1. We assigned the case-weighted mean of the observed CMI of hospitals within a

stratification cell to be the prediction for all hospitals in the validation sample that fell into that stratification cell.

Model 3. Predict CMI by stratifying the sample with respect to case-level characteristics.

We stratified all cases in the fitting sample on the basis of predicted RIC, transfer status (transfer to hospital, transfer to SNF, or not a transfer), and an indicator of whether the observed length of stay of a case was greater than the mean length of stay for cases in the RIC category corresponding to the predicted RIC. The mean length of stay of all RICs was used for cases without a predicted RIC. We assigned each case in the validation sample a prediction equal to the mean weight of the fitting sample cases that fell into the corresponding stratification cell. After assigning predictions to the cases in the validation sample, we computed the predictions of CMI for hospitals in the validation sample as the average of the case weight predictions.

In addition, we predicted CMI using a related model:

Model 3a. Predict CMI by stratifying the sample with respect to the case-level characteristics described above; include comorbidity tier.

Model 4. Predict case mix with a linear regression model.

We regressed case-level weight on the following case-level predictor variables: predicted comorbidity (yes/no/unknown); predicted RIC (including unknown); transfer to a SNF (yes/no); length of stay (in days); and transfer status (yes/no). To account for the interaction between LOS and transfer status, a linear spline term for the LOS and transfer interaction was included in the model. Prior analysis of the case-level data indicated that the optimal cutoff for a 1-knot linear spline term for the LOS and transfer interaction term is 14 days. We set LOS and transfer interaction equal to 0 for those who had less than a 14-day LOS, and equal to LOS otherwise.

ESTIMATION OF THE CONVERSION FACTOR

To understand how well we can predict CMI, it is important to understand how any error from prediction will translate in terms of the conversion factor.

To derive the payment per hospital, we first had to derive a conversion factor based on all the CMI predictions for all hospitals in

the sample. The goal of this analysis was to assess our CMI prediction method by comparing the conversion factor resulting from the predictions of CMI to the one resulting from using the observed CMI values. If we calculate x as the sum across hospitals of what IRF payments would be if the conversion factor were 1, then the budget- neutral conversion factor is the sum of payments made to hospitals under TEFRA divided by x. If any other conversion factor is used to compute non-budget-neutral IRF PPS payments, then the budget-neutral conversion factor is that conversion factor multiplied by the ratio of TEFRA payments to the non-budget-neutral IRF PPS payments.

Four main steps must be completed to compute the conversion factor and to evaluate prediction error. We detail each step below.

Step 1. Compute total payments under the TEFRA system.

To obtain the average pay per discharge under TEFRA, compute the total TEFRA payment per hospital as the sum of operating and capital payments; then divide by the number of discharges in the OACT cost report for that hospital. To account for the small number of discharges bundled into cases under the IRF PPS, multiply the TEFRA payment per discharge by the ratio of discharges to cases in the MEDPAR to get TEFRA payment per sample case. Finally, multiply by the number of sample cases to obtain the total pay for these cases under TEFRA at each hospital. Add the total TEFRA payments across hospitals; the result is the numerator of the budget neutrality calculation. Obtain the "weighted (by cases) average TEFRA pay" by dividing this numerator by the total number of sample cases at those hospitals.

Step 2. Compute the payment under the new IRF PPS without accounting for budget neutrality.

To calculate the IRF PPS payment, the weighted average pay under TEFRA (derived from Step 1) is used as an initial estimate of the conversion factor. Then total IRF PPS payment is computed as the product of this initial conversion factor, CMI, number of sample bundles, an outlier adjustment (1 + 0.03/0.97, which reflects outlier payments equal to 3 percent of total payments), and facility-level adjustment factors. The facility-level adjustment factors included in our IRF PPS payment are wage index, rural status, and disproportionate share (low-income) percentage. The magnitudes of these facility-level adjustments are

derived from Section 7, except that we used the updated percentage of labor related share. So the adjustments used here are:

Wage index adjustment: (0.27605 + 0.72395 * (wage index))

Rural adjustment: (1.0 + 0.1914 * (rural status (1=rural; 0=not rural)).

Disproportionate share adjustment: ((1.0 + DSH) ** 0.4838).

The result is the new payment without budget neutrality for each hospital. To obtain the denominator of the conversion factor ratio, add all of the new payment values for the group of hospitals.

Step 3. Compute the budget-neutral conversion factor.

Divide the total payment under TEFRA (from Step 1) by the sum of the new payment without budget neutrality (from Step 2). Multiply the resultant ratio by the initial estimate of the conversion factor (the average TEFRA payment from Step 1) to get the budget-neutral conversion factor.

Step 4. Evaluate the conversion factor prediction error.

Ideally, we would like to test our models on data that are separate from the data used to create our models. However, we also wanted to use all available data for developing a model to predict CMI. To address both of these concerns, we used cross-validation to obtain an estimate of the out-of-sample prediction error (Efron and Tibshirani, 1993). We split our hospital sample into ten roughly equal parts so that there were approximately 65 hospitals per part. We then combined nine of the ten parts to be a "fitting sample" on which we fit our model to the data, and held out the tenth part to be the "validation sample" from which to compute a prediction error estimate. We repeated this ten times, each time holding out a different tenth of the data as the validation sample, and averaged the ten prediction errors at the end of the analysis to obtain the cross-validated prediction error estimate.

We evaluated the error by comparing the conversion factors computed using the observed and predicted CMIs, where predicted CMI was computed as described above (i.e., in Models 1-4). In addition to these predictions, we also examined a "Model 0," for which CMI was not predicted at all. For Model 0, the budget-neutral conversion factor was computed for the fitting sample and applied to the validation sample. To assess the error of this method, we compared the new payment with budget

neutrality that would be obtained by using the observed CMI values to the payment under Model 0. (Model 0 is essentially the method described in the NPRM.)

RESULTS

Table 9.3 shows the out-of-sample prediction error estimates obtained from four separate cross-validations for Models 0-4. Column 1 lists the models for each of the four runs; column 2 shows the average observed conversion factor (CF) in the runs. The average observed CF is the average of each of the conversion factors for the validation sample in each of the ten partitions of the data and thus varies from run to run. Column 3 shows a similar average for the predicted CF. The root mean square error (RMSE) in column 4 is a measure of the typical size of the error in predicting the conversion factor for the out-of-sample hospitals. Column 5 expresses the RMSE as a percentage of the actual conversion factor, and column 6 expresses the RMSE as a percentage of the RMSE for Model 0.

Although the numbers vary slightly across the four runs, several conclusions can be firmly drawn from these data. First, we found that Model 1 (and indeed all the models) provides a much more accurate prediction of the conversion factor for out-of-sample hospitals than does Model 0. The RMSE is reduced by about 40 percent on average. Thus we are better off using our data to estimate the CMI at hospitals and then using the data from the non-sample hospitals on TEFRA payment and facility adjustment to estimate a conversion factor.

Model 1 assigns the same CMI to all out-of-sample hospitals. Model 2 assigns an estimate of CMI for each hospital group defined by whether it is freestanding, a unit, or rural, and two wage index categories. These characteristics only modestly affect CMI, and the result is effectively no improvement in RMSE compared to Model 1.

Table 9.3

**Prediction Error Estimates in Terms of the Budget-Neutral
Conversion Factor**

Model	Average Observed CF ($)	Average Predicted CF ($)	RMSE ($)	% Error	RMSE as % of Model 0
Run 1					
0	11,875.20	11,854.48	386.94	3.26	100.00
1	11,875.20	11,889.10	216.17	1.82	55.87
2	11,875.20	11,892.37	213.01	1.79	55.05
3	11,875.20	11,875.20	184.83	1.56	47.77
3a	11,875.20	11,874.59	180.09	1.52	46.54
4	11,875.20	11,866.90	177.83	1.50	45.96
Run 2					
0	11,851.66	11,854.18	433.27	3.66	100.00
1	11,851.66	11,881.08	180.30	1.52	41.61
2	11,851.66	11,884.50	180.66	1.52	41.70
3	11,851.66	11,856.24	112.05	0.95	25.8
3a	11,851.66	11,854.04	99.60	0.84	22.99
4	11,851.66	11,850.19	134.43	1.13	31.03
Run 3					
0	11,872.37	11,854.39	330.99	2.79	100.00
1	11,872.37	11,897.29	270.59	2.28	81.75
2	11,872.37	11,901.46	274.62	2.31	82.97
3	11,872.37	11,894.38	218.29	1.84	65.95
3a	11,872.37	11,891.45	203.47	1.71	61.47
4	11,872.37	11,892.86	230.55	1.94	69.65
Run 4					
0	11,825.52	11,853.89	419.02	3.54	100.00
1	11,825.52	11,855.42	285.18	2.41	68.06
2	11,825.52	11,856.46	279.04	2.36	66.59
3	11,825.52	11,837.21	241.61	2.04	57.66
3a	11,825.52	11,836.12	220.49	1.86	52.62
4	11,825.52	11,834.10	259.69	2.20	61.98
Average					
0	11,856.19	11,854.24	392.56	3.31	100.00
1	11,856.19	11,880.72	238.06	2.01	60.64
2	11,856.19	11,883.70	236.83	2.00	60.33
3	11,856.19	11,865.76	189.20	1.60	48.20
3a	11,856.19	11,864.05	175.91	1.48	44.81
4	11,856.19	11,861.01	200.63	1.69	51.11

Models 3, 3a, and 4 used case characteristics to predict the hospital CMIs. Models 3, 3a, and 4 resulted in improvements over Models 1 and 2 in all four runs. The size of the improvement varied by run. Compared to Model 1, the improvement due to Model 3 varied from about 15 percent in runs 1 and 4 to 38 percent in run 2; Model 3a varied from 16 percent (run 1) to 45 percent (run 2); whereas the improvement due to

Model 4 varied from 9 percent in run 4 to 25 percent in run 1. Compared to Model 0, the improvement due to Model 3 ranged from 34 percent in run 3 to almost 74 percent in run 2. Comparable improvements were seen with Model 3a, and the improvement in Model 4 over Model 0 ranged from 30 percent in run 3 to 69 percent in run 2. On average, the RMSE under Model 3 was less than half of that under Model 0. Thus, these case-level data predicted CMI effectively in all runs, and could have a large effect when the conversion factor of sample hospitals differs from that of out-of-sample hospitals. Model 4 performed slightly better than Models 3 and 3a in run 1, but worse in the other runs. However, Model 4 was substantially better than Models 1 and 2 as well as Model 0.

In summary: Model 0 consistently had the largest prediction error (RMSE), which reflects substantial uncertainty associated with this prediction and thus makes it a poor choice for computing the conversion factor. Model 1 (case-weighted mean CMI prediction) and Model 2 have similar RMSEs, substantially better than that of Model 0 and worse than those of Models 3 and 4. The RMSE generally decreases as one adds more information to the model in the form of case or hospital characteristics, thus helping to explain variation that otherwise would appear as prediction error. As is clear from the RMSE, the addition of detailed, case-level characteristics (including predicted RIC and predicted comorbidity status) in Models 3 and 4 improves the prediction of CMI in terms of conversion factor, as compared to Models 0-2.

Model 3a performed best in three of the four runs. In the remaining run, run 1, it performed only slightly worse than Model 4. Thus it appears that Model 3a produces the most accurate predictions of the conversion factor.

IMPLICATIONS FOR A NATIONAL CONVERSION FACTOR

We wanted to estimate the national conversion factor that would produce budget neutrality if the cases in FY 2002 at each hospital were to look like the cases at that hospital in CY 1999 and average TEFRA payments at each hospital conformed to the OACT estimates. The analysis in the last subsection showed that Model 0, the methodology proposed in the NPRM as the best available at that time, results in slightly more than a 3 percent error in estimating the conversion factor that would

produce budget neutrality at a randomly selected 10 percent of the sample hospitals. In using this methodology to estimate the FY 2002 conversion factor, we actually know the case mix for about 60 percent of the cases. Thus the error in the conversion factor affects only about 40 percent of the dollars, and thus the average error is reduced to 1.2 percent (= 100 * 0.03 * 0.40). If the hospitals in our sample are a randomly selected subset of the universe and the cases in our matched FIM-MEDPAR data are randomly selected, the expected conversion factor error will be further reduced by the law of large numbers (the 3 percent refers to a sample of only about 60 hospitals, whereas the hospitals outside our entire sample number slightly under 500).

Our sample, however, is not randomly selected in at least two ways. First, hospitals choose to join or not join UDSmr, and our sample also contains all (or almost all) of the 1999 cases at hospitals owned by a single large corporation. Thus the sample is not a random selection of hospitals and indeed underrepresents certain types of hospitals, including units and rural hospitals.

Second, the hospitals choose to fill out the FIM for each case and send it to UDSmr or to corporate headquarters. At UDSmr, these data are used to prepare profiles of hospital case mix to compare against those of the UDSmr national database. We obtained FIM data for about 90 percent of MEDPAR discharges at hospitals that participated in the central database throughout 1999 (Table 2.2).[3] Although this is a very large proportion, it is lower than can be explained by expected error rates in the demographic and date data used in the match. Insofar as the other 10 percent of the cases are due to random errors in the match, they are indeed randomly selected. However, if hospitals choose to selectively withhold particular kinds of cases from the central database, our selection cannot be random. If the withheld cases have an average case weight that differs from that of the included cases, it could distort the conversion factor.

The results in the last subsection also showed that our other models could account for some of the effect on the conversion factor of the non-representative nature of our sample. In this subsection, we show

---

[3] Further details are in Relles and Carter (forthcoming).

how we used each model to estimate a national conversion factor for FY 2002. We did so by combining the actual data for the sample cases with predictions from the model for other hospitals and/or cases.

As shown in the first row of Table 9.4, Model 0 produces an estimated conversion factor of $11,854. This estimate is the budget-neutral conversion factor that makes IRF PPS payment equal to FY 2002 TEFRA payment for the 233,043 bundled cases in our matched FIM-MEDPAR CY 1999 database from hospitals in the OACT file.

The remaining models begin with the same 233,043 cases from the same 659 hospitals. They then use these data in two ways: first, to predict the case weight of additional FY 1999 rehabilitation cases, and second, to combine with these predictions in order to estimate a national conversion factor.

Table 9.4

Estimates of the Conversion Factor from Various Models and Samples

Model	Hospitals	Cases	Conversion Factor ($)	% Increase from Model 0
Model 0	659	233,043	11,854	
Model 1	1,024	316,813	11,830	-0.20
Model 2	1,024	316,813	11,847	-0.06
Model 3-P1	659	264,019	11,952	0.83
Model 3-P2	1,024	347,782	11,994	1.18
Model 3a-P1	659	264,011	11,948	0.80
Model 3a-P2	1,024	347,757	11,977	1.04
Model 4-P1	659	264,039	11,985	1.11
Model 4-P2	1,024	347,809	12,025	1.45

Model 1 calculates the case-weighted mean CMI at the sample hospitals. It calculates the conversion factor assuming that the CMI at each non-sample hospital is this mean CMI and that the CMI at each sample hospital is that from the sample discharges. Thus, Model 1 takes into account how the non-sample hospitals differ in their facility adjustments and TEFRA payments. Model 2 adds how they differ in CMI due to the hospital characteristics in the stratification discussed in Section 4.

As shown in Table 9.4, the CFs from Models 0, 1, and 2 are all very similar. Thus we believe that the selection of hospitals into our sample

has little effect on the CF, at least insofar as they are non-representative on TEFRA payment, wage index, rural status, and freestanding or unit.

Models 3 and 3a add information about the case characteristics of the non-sample cases. As we saw in Section 6, these models are much better predictors of case mix than is Model 2, which is built only on hospital characteristics. Models 3 and 3a (and Model 4) can be applied in two ways. The rows labeled Model 3-P1 and 3a-P1 show our prediction of case weight for each MEDPAR case that is not part of the matched FIM-MEDPAR database but is discharged from a sample hospital. In Models 3-P2 and 3a-P2, we used the same model to predict the case weight for all the MEDPAR cases that are not in our database.

Surprisingly, applying Models 3 and 3a only to non-matched cases from sample hospitals resulted in a 0.8 percent increase in the CF compared to the Model 0 CF. Applying the model to all non-sample cases resulted in a further increase--1.0 percent greater than Model 0 for Model 3a-P2 and 1.2 percent greater for Model 3-P2. Interestingly, the conversion factor from Model 4 applied to the universe is similar to those of Models 3 and 3a.

Eighty-five percent of the cases used in Model 3-P1 were also used in Model 0. Thus the case mix of the 15 percent of non-sample cases from sample hospitals must be substantially lower than that of the matched sample. We have contrasted these cases in a variety of ways and conclude that this is a real effect of real differences in case weights, not a statistical artifact.

Table 9.5 shows the distribution of cases in the 1998-1999 sample by each value of predicted RIC. The second column shows the percentage of these cases for which we have FIM data. So, for instance, there were 94,158 cases with a predicted RIC of stroke (this group is easy to characterize because it consists of cases where the rehabilitation immediately followed an acute stay for stroke care). Of these we had FIM

Table 9.5

**Percentage of Cases from Sample Hospitals That Are Not in FIM Database,
by Predicted RIC (combined 1998 and 1999 data)**

Predicted RIC	Number of Cases	Percentage Not in Sample
1	94,158	13.2
2	5,971	14.6
3	6,599	13.1
4	621	11.9
5	22,884	13.7
6	13,677	15.1
7	69,036	13.7
8	104,246	12.1
9	14,706	13.9
10	19,842	14.9
11	2,626	18.1
12	1,511	17.5
13	1,139	17.5
14	29,393	14.7
15	6,561	13.7
16	892	16.1
17	2,075	15.2
18	589	18.2
19	1,051	15.3
20	54,959	13.6
21	457	32.4
Atypical short stay	17,236	37.1
Deaths	2,934	22.4
Missing	102,645	17.8
Total	575,808	15.0

data for all but 13.2 percent--only slightly lower than the 15.0 percent
overall for which we are missing FIM data.[4] For most values of predicted
RIC, the proportion of non-sample cases is within a few percent of 15.

---

[4] The 15 percent missing includes hospitals that were in our
database for only part of the year and thus contributed only a small
fraction of their total cases to the database. The 90 percent match rate
mentioned above and in Table 2.2 is restricted to hospitals that were in
the database throughout the year.

But there are two exceptions--the small burn RIC 21[5] and the atypical short-stay cases. We are missing FIM data for fully 37 percent of the atypical short-stay cases. Although they constitute a relatively small percentage (3 percent) of total cases, they are disproportionately missing from the sample. Because of their very low weight (roughly 16 percent of average), they have a measurable impact on the estimated CMI at these hospitals. Indeed, alone they account for much of the 0.8 percent increase in the CF in Model 3a compared to Model 0.

We believe these data are real in the sense that they represent real cases--included in the TEFRA discharge counts and in the average payment and cost statistics. Thus we believe that Model 3a (either for the sample or for the universe) should be used to calculate the conversion factor in the final rule.

CMS chose to use the predictions from Model 3a on the entire population of cases for which CMG weights could be estimated as the best estimate of what the budget-neutral conversion factor would be if the cases in FY 2002 were exactly like the cases in CY 1999 and there were no behavioral response. This conversion factor is 1.04 percent higher than the conversion factor using the methodology of the NPRM. The CMS Actuary then used his judgment and experience to develop the behavioral offset that reduced the conversion factor by 1.16 percent.

---

[5] The underrepresentation of the burn cases in our sample is almost surely due to the fact that some of these cases are not rehabilitation cases and will not be paid under the PPS. In our sample this is true of a very small sample of cases--roughly 70. Although their average costs are high, the total amount of funds involved cannot affect the CF in any measurable way.

10. DEVELOPING A MONITORING PLAN AND SYSTEM FOR THE IRF PPS

A major focus of our ongoing work will be the development of a system to monitor the impact and performance of the IRF PPS. This system will serve two primary purposes. First, it will track patient access to IRF care and changes in the costs, quality, and outcomes of IRF care.[1] Second, the system will monitor changes in the care delivered across post-acute care settings. This is important because the financial incentives created by the IRF PPS may affect the number and mix of patients using other types of post-acute care, and because changes in the payment systems for other types of care may affect the patients admitted to IRFs.

The monitoring system will be used to assess the positive and negative effects of the IRF PPS and to determine if the PPS is meeting its intended goals. The main goals of the IRF PPS are

- To ensure beneficiary access to high-quality, medically necessary, inpatient rehabilitation care
- To promote efficiency in the delivery of IRF care
- To compensate providers fairly
- To avoid unintended consequences and perverse incentives.

Monitoring whether the system is meeting these goals will include identifying regional and national trends in IRF care and identifying unusual hospitals. However, no data that identify, or could be used to identify, individual hospitals will ever be publicly released. The monitoring system will also be used to identify areas where the IRF PPS needs refinement and to assess the effects of the IRF PPS on other types of post-acute care.

The monitoring system will be designed with input from the TEP and from CMS and will draw on analyses conducted during phase I. Baseline measures of cost, access, and quality will be developed over the next year. Post-implementation data will be analyzed and evaluated starting

---

[1] Our ability to measure quality and outcomes will be limited, as described below.

in fall 2002. Below we describe our basic approach to the development of the monitoring plan and outline the measures that we believe should be incorporated into the monitoring system.

QUESTIONS THAT WILL GUIDE THE DEVELOPMENT OF THE MONITORING SYSTEM

Our goal is to answer a series of questions about the direct effects of the IRF PPS. These questions include the following:

- What are the effects of the IRF PPS on beneficiary access to rehabilitation care, the mix of services delivered to patients, the costs of and payments for that care, and the outcomes of that care?

- What are the effects of the IRF PPS on individual IRFs, the market(s) for rehabilitation care, and the market(s) for post-acute care?

- What are the causes and outcomes of the wide variation in the use of post-acute care across markets? How has post-acute care delivery changed following implementation of the IRF PPS?

- What effect does the IRF PPS have on the costs of the Medicare program as a whole?

We will also seek to answer questions about how well the components of the IRF prospective payment system are performing:

- How well do the case mix weights predict costs post-implementation?

- Are the comorbidities incorporated into the payment system still important drivers of costs?

- Are the outlier targets appropriate?

- Do the facility adjustments correctly reflect the extra costs of operating in rural areas and serving low-income patients?

- Is the system budget neutral?

These questions were discussed in more detail in previous sections of this report.

In addition, we will keep in mind three other broad questions related to Medicare's role as a prudent purchaser on behalf of beneficiaries and taxpayers. We will not be able to answer these questions directly, but we will use them to guide the development of the

monitoring system. We also hope that the monitoring system will shed some light on the likely answers to these questions.

First, what are the most appropriate (i.e., effective and efficient) levels of and settings for care of different types of rehabilitation patients? How do the effectiveness and efficiency of care vary by patients' diagnoses, complexity (including impairments, levels of functioning, and comorbidities), and circumstances, including the availability of a caregiver where they reside?

Second, to what extent are patients receiving rehabilitation care that is well matched to their diagnoses, complexity, and circumstances? How has that changed with the introduction of post-acute care PPSs, including the IRF PPS? How much variation is there in the care delivered within post-acute care settings?

Third, perhaps the broadest question we will keep in mind is, what is the marginal benefit of rehabilitation therapy? What gains in functioning do patients achieve that they would not have achieved or would not have achieved as quickly without rehabilitation therapy? How do the costs and benefits of rehabilitation therapy compare to the costs and benefits of other medical or social services that Medicare does or could provide?

These efforts will be constrained by the current state of the science in rehabilitation medicine. Only a few studies have examined patient outcomes in post-acute care settings (see Keith, Wilson, and Gutierrez, 1995; Liu et al., 1999; MedPAC, 2001). What little research there is in this area has focused on outcome differences across post-acute care settings, rather than focusing specifically on measuring outcomes for inpatient rehabilitation. In comparing post-acute care settings, Kane et al. (1997a) found that rehabilitation facilities, relative to SNFs and home health care, demonstrated the greatest potential for functional improvement for stroke patients who had high ADL dependency scores at discharge. In a related study, Kane et al. (1997b) compared patients' actual discharge locations with those that would have produced "optimal outcomes." They found that a very low percentage of patients (23 to 50 percent) were placed in the post-acute care setting that would provide the most functional improvement. In

studying whether outcomes and costs differ for elderly patients admitted
to rehabilitation hospitals, sub-acute care nursing homes, and
traditional nursing homes, Kramer et al. (1997) concluded that the more
comprehensive therapy services provided by rehabilitation facilities can
lead to better functional recovery and community placement for stroke
patients, but with the effect of higher costs. These higher costs did
not appear to benefit hip fracture patients. Keith, Wilson, and
Gutierrez, (1995) compared acute and sub-acute rehabilitation for stroke
patients and found that acute rehabilitation patients showed
substantially greater gains in functional measures, but at more than
double the cost per successful discharge than the less intensive sub-
acute setting. Other studies, such as Liu et al. (1999), vary in their
findings. Liu et al. (1999) found that rehabilitation patients had
lower death rates and rehospitalizations than did those treated in sub-
acute SNFs. Most outcomes were not significantly different for patients
treated in sub-acute SNFs and inpatient rehabilitation facilities after
controlling for patient characteristics.

Munin et al. (1998) performed the single major study focusing
solely on inpatient rehabilitation facilities. The researchers compared
the effects of time lapse between surgery and admission to
rehabilitation. Patients undergoing elective hip or knee arthroplasty
and who began inpatient rehabilitation three days after surgery attained
short-term functional milestones in fewer days using fewer total
resources compared to patients who began rehabilitation seven days after
surgery.

MONITORING THE EFFECTS ON IRF CARE

The first goal of the monitoring system will be to develop
indicators that measure the care delivered by IRFs before and after the
implementation of the PPS. These will include measures of payments for IRF
care, measures of beneficiary access to IRF care, measures of the "health"
of rehabilitation facilities, and measures of the quality and outcomes of
care delivered by IRFs. In developing this set of measures we will keep in
mind the key concerns of policymakers about incentives under the IRF PPS.
These concerns are outlined in Figure 10.1.

```
• Access to care:
  - Complex cases might experience reduced access to IRF care.
  - Less-complex cases might be admitted unnecessarily.
• Underutilization of care:
  - Beneficiaries might receive inappropriately lower intensities
    of care and/or shorter lengths of stay in IRFs.
  - IRFs might discharge patients prematurely.
• Upcoding:
  - IRFs might code beneficiaries into higher weight CMGs and tiers.
```

Figure 10.1--Key Concerns About Incentives Under the IRF PPS

All these concerns relate to the new financial incentives that IRFs will be faced with. Previously, IRF reimbursement depended on a per-case target amount based on the facility's historical costs and on the facility's actual cost per case. The Balanced Budget Act of 1997 introduced maximum payment limits, or caps, that affected some facilities. Now that facilities receive a fixed payment that varies by patient severity, they have increased incentives to contain costs since their profit will be equal to the difference between the fixed payment and their actual costs. This means that patients whose cases are more complex than the average case in their payment group, or for whom there is more uncertainty about their care costs, may experience reduced access to care. Conversely, patients whose cases are less complex or more predictable may be admitted at higher rates if facilities believe they are more profitable. (The reverse, however, could happen in facilities that are at their payment limit. They may be more able to admit high-cost patients under the IRF PPS.) In addition, IRFs may change their coding practices. Previously, facilities had no incentive to code patient diagnoses accurately or completely. Because comorbidities now bring extra payments, facilities are likely to step up their efforts to code them. Facilities can also try to increase their margins by incurring fewer costs in treating patients. Indeed, greater incentives for efficiency are a goal of the IRF PPS. However, facilities might attempt to reduce patients' lengths of stay to levels that could

affect them adversely. They could also try to reduce length of stay by discharging patients earlier and sending them home with home health care or to nursing facilities to finish their recovery. This strategy would work against the intent of the payment system to the extent that it would increase total Medicare payments.

Access-to-Care Indicators for IRFs

We envision several approaches to examining whether the outcomes of the IRF PPS justify these concerns. We will begin by developing a series of indicators that measure use of and access to IRF care over time in order to gauge the effects of the IRF PPS. First, we will establish baseline rates and measure underlying trends in the characteristics of patients receiving care at IRFs. In particular, we will look at the fraction of acute care patients discharged to IRFs by DRG. Our goal is to determine how this fraction varies over time and across markets. We will also track trends in the CMG mix of IRF patients and in comorbidity rates among IRF patients by facility and facility type (e.g., hospitals that were close to or at their payment limit versus those that were not). We will model predicted RICs and comorbidities based on data about patients' acute stay and compare them to the actual RICS and comorbidity tiers billed by IRFs to gauge the extent of coding change. We will also look at changes in the characteristics of IRF patients that are not related to the payment system but might affect IRF care, such as the proportion of patients living at home with family members prior to their hospitalization. Finally, we will evaluate the cost implications of these trends for facilities, for the Medicare program, and for beneficiaries' out-of-pocket costs.

Monitoring Trends in Intensity and Outcomes of IRF Care

We will measure and track changes in the intensity of care delivered by IRFs and in the outcomes of that care. Using FIM admission and discharge scores and information from the new IRF PAI,[2] we will

---

[2] Facilities will provide FIM and functional modifier data that describe patients during the three days after admission and at discharge.

examine the cost data submitted by facilities in order to track changes in facility costs per case and the intensity of care delivered. In addition, we will track length of stay in IRFs, controlling for changes in the mix of patients seen by those facilities.

Post-IRF care outcomes will also be studied. We will establish baselines and monitor such changes in health status measures as mortality rates and readmission rates to acute care hospitals. We will also monitor trends in discharge settings after IRF care. Specifically, we will track national and regional changes in the use of other types of post-acute care after IRF discharge, including the number and intensity of services and the amount of Medicare payments for those services. In conducting all of these analyses we will control for changes in the types of patients seen at IRFs, using information from our analyses of care use described above. (This will be achieved by modeling the decision to use post-inpatient-rehabilitation care: The model will include patient characteristics, hospital characteristics, measures of the financial pressure on the facility produced by the IRF PPS, etc. Preliminary work has revealed that post-IRF use of home health care varies significantly by patient age, impairment, comorbidities, region, and other factors.)

MONITORING EFFECTS ACROSS POST-ACUTE CARE PROVIDERS

Our monitoring system will also be designed to monitor changes in post-acute care delivered across the spectrum of post-acute care providers, which include home health agencies, skilled nursing facilities, outpatient departments, Comprehensive Outpatient Rehabilitation Facilities (CORFs), long-term care hospitals, and hospices. This is important because the change in financial incentives caused by the IRF PPS might have "spillover" effects on other types of post-acute care. For example, if IRFs were to discharge their patients earlier and send them home with home health care at higher rates, the mix of patients using home health care would change. The IRF PPS should function in conjunction with other payment systems to create incentives for beneficiaries to receive services in the most appropriate post-acute care setting. As with the indicators for monitoring IRF care, we will

develop a set of indicators for post-acute care that are in line with
the key concerns of policymakers about the effects of prospective
payment on post-acute care. These are outlined in Figure 10.2.

- Access to care:
  - Prospective payment systems for post-acute care might limit
    types of care available for complex cases.
  - Financial considerations might override clinical
    considerations in determining the choice of post-acute care
    site.
- Utilization of care:
  - Patients might be shuttled among sites of care to increase
    Medicare payments.

**Figure 10.2--Key Concerns About Effects of PPSs on
Post-Acute Care**

Other post-acute care prospective payment systems, like the IRF
PPS, give facilities incentives to contain costs. They may, therefore,
limit access to care for patients who are expected to incur more costs
than the fixed payment the facility will receive from Medicare for their
treatment. Table 10.1 describes some of the key payment system
attributes that will likely affect patient use of care. Because each of
the post-acute PPSs has different units of payment and levels of payment
for different but overlapping categories of patients, however, one type
of provider might be relatively well compensated for a certain patient
whereas another is not. For example, it could be that a hip fracture
patient with multiple ADL limitations, good cognitive functioning, and
comorbid breast cancer is a profitable patient for IRFs and home health
care agencies but an unprofitable one for skilled nursing facilities. In
this case, financial considerations could override a clinical decision
that a SNF was the most appropriate post-acute care site. In addition,
if that patient were admitted to an IRF or a SNF, the facility could

327 -

Table 10.1

Post-Acute Care Prospective Payment Systems

Characteristic	Skilled Nursing Facility	Home Health Agency	Inpatient Rehabilitation	Long-Term Hospital	Outpatient Department
Unit of payment	Per diem	60-day episode (unlimited episodes)	Per discharge	Per discharge	Per service
Patient classification instrument	Minimum Data Set 2.0	Outcome and Assessment Information Set	IRF Patient Assessment Instrument (including FIM items)	UB-92 (and cost report settlements)	HCFA-1500
Classification system	RUG-III	HHRGs	CMGs	Refined DRGs	APCs
Adjustments	Wage	5% outlier; transfer; low util.; SCIC; wage	3% outlier; transfer; rural; DSH; wage	Bonus/Relief	2.5% outlier; wage

have an incentive to discharge her as soon as possible--the SNF because each day of providing expensive cancer drugs causes it a loss, and the IRF because the facility receives no extra payments for extra days of care. Thus, the different post-acute prospective payment systems may create complex sets of financial incentives. This could be particularly problematic since there is currently little clinical consensus about, and wide variation in, where patients are sent for post-acute care.

Post-Acute Wide Effects of New Payment Systems

To develop specific hypotheses about the wider effects of the IRF PPS, we will analyze each of the post-acute payment systems. We will create a comprehensive database that tracks the payment systems used for each type of post-acute care from 1996 through the end of our data collection period in 2003. Figure 10.3 presents the timeline for the implementation of post-acute care payment systems from 1997 onward. We will develop models of the likely payments that would be associated with patients depending on the setting to which they were discharged. Using this information, we will develop and test hypotheses about how changes in the number and types of patients receiving care in each post-acute setting will respond to payment system changes.

Monitoring Trends in Use of Post-Acute Care After Implementation of the IRF PPS

We will develop indicators and track trends in the fraction of acute care patients receiving care at each type of post-acute care setting (for selected DRGs) by market and for the nation as a whole. We know that there is currently wide variation in the use of post-acute care across the country and also wide variation in the content of post-acute care within at least some settings (e.g., the intensity of rehabilitation therapy offered in SNFs). We will use market characteristics (capacity of IRFs and SNFs, hospital wage rates, HMO market share, etc.) to try to understand the causes of this variation in the pre-IRF PPS market and why markets differ in their response to the IRF PPS.

We will assess the cost implications of these differences for IRFs, for post-acute care providers generally, for beneficiaries, and for the Medicare program as a whole. IRF costs will be examined using IRFs' cost
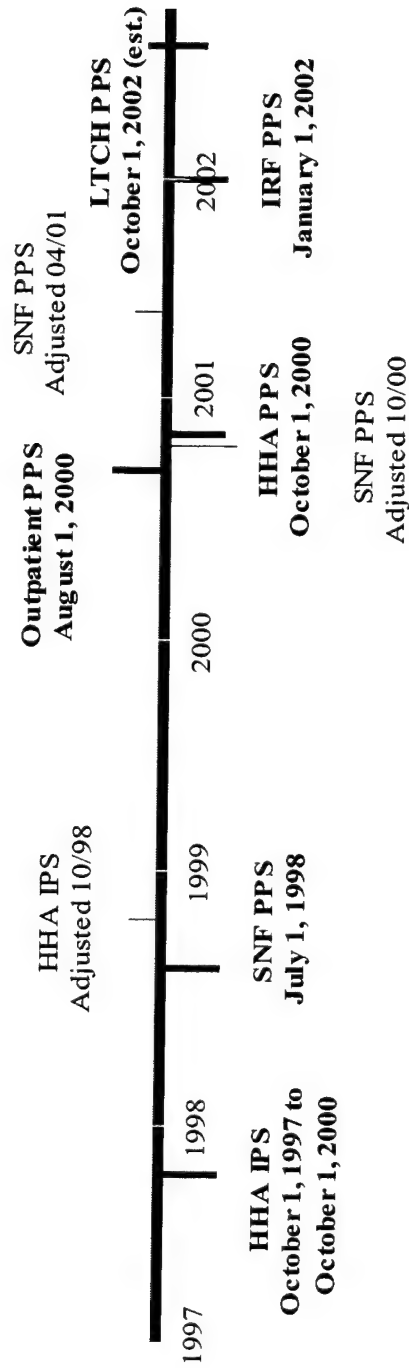
Figure 10.3--PPS Implementation Timeline

reports. We will not be able to estimate changes in costs of other post-acute care providers directly, but we will examine how their case mix might change by using information from patients' acute care stays. Total Medicare program payments and payments to beneficiaries for "episodes" of post-acute care will be examined using Medicare claims data.

Monitoring Trends in Outcomes of Post-Acute Care

We will also measure improvements in health status associated with care in each post-acute care setting both at baseline and over time. The measures of health status that we intend to use include mortality rates, acute hospital readmission rates, and length of stay in formal care. These measures will be tracked across markets and over time to see if the rate of use of types of care is associated with better outcomes. Because we realize that these measures of outcomes are very crude, we will seek to obtain data from other sources that would allow us to measure outcomes, such as institutional status, at fixed intervals following an admission (e.g., three months). We will also explore what would be required to undertake the collection of such outcome data.

Sampling Strategies for Post-Acute Care Monitoring

Our studies of post-acute care will likely be based on a sample of post-acute care users--unlike our work on IRF care, which will encompass the entire population of Medicare IRF users. We plan to pull a sample of DRGs representative of post-acute care users. In doing so, we will take steps to make sure that we sample clusters of DRGs, when necessary, to include clinically related categories. For example, we will sample the DRGs for "Back and Neck Procedures" both with and without complications. In preliminary work, we have identified stroke, major joint replacement, hip fracture, and spinal fusion as candidates--although we expect to use a greatly expanded list. In addition, we may select certain DRGs for in-depth studies of selected patient types. These could include DRGs for which hospital admissions are not discretionary, conditions which are considered particularly responsive to rehabilitation treatment, and conditions about which there is heightened concern about potential quality problems (e.g., where the patient is cognitively impaired).

We may also draw a random sample of post-acute care users--
potentially with an oversample of selected markets so that we have the
ability to capture market effects.

ANALYTIC ISSUES

In each of these analyses, we will have to account carefully for
differences in patient severity and circumstances so that we correctly
attribute changes to the IRF PPS rather than to other factors. This
means that we will have to model the factors that affect sites, costs,
and outcomes of care, and we will have to undertake analyses that
account for the fact that patient severity is not entirely observable.
We will also have to exercise care in choosing appropriate baselines
given the level of change in the market and evolving medical technology.
Finally, we will have to use rigorous methods to determine "problem"
levels from statistical and policy perspectives.

We will address these analytic issues by identifying appropriate
sources of comparison. Generally speaking, the monitoring and assessment
of the IRF PPS will be conducted by comparing costs, quality, access,
outcomes, etc., to standards or averages. Our key sources of comparison
will be pre- and post-PPS implementation. We will, however, also probe
how pre- and post-implementation changes vary within and between RICs,
IRFs, and markets (i.e., geographic areas). By exploiting natural
variation in these areas, we will be able to shed some light on our
larger questions about whether the changes produced by the IRF PPS are
beneficial for patients, facilities, and the Medicare program.

To do this, we will draw on the work of researchers who have
examined the variability in use of post-acute care. For example, a
number of researchers have explored cross-sectional regional variation
in the use of various types of post-acute care. Kenney and Dubay (1992)
found that MSAs with fewer nursing home beds per enrollee, higher
hospital discharge rates, shorter mean LOSs, higher ceilings for HHA
payments, and more home health agencies per enrollee used more home
health care. They also found that the number of home health visits was
higher in areas with more agencies and where more agencies were
proprietary. Lee, Huber, and Stason (1997) analyzed cost and service use
data for a 20-percent random sample of Medicare beneficiaries

hospitalized for stroke in 1991. Market- and patient-level variables in their models explained approximately one-half of the variation across MSAs in post-stroke rehabilitation practice patterns. Liu, Wissoker, and Rimes (1998) predicted post-acute use of SNF and HHA using data from the 1991 Medicare Current Beneficiary Survey. They found that patient characteristics, including age, marital status, and number of ADL limitations, were predictors of any use of SNF or HHA care. The level of availability of Medicaid home health care was also a predictor of use. For the use of SNF care versus HHA care, they found that a greater number of patient characteristics came into play, including the beneficiaries' place of residence. Cohen and Tumlinson (1997) found that use of Medicare home health by MSA varies inversely with the amount that Medicaid spends on home health. They also found an inverse relationship between Medicare home health use and the number and use of nursing facilities.

We will also explore whether it is possible to track changes in the Medicare versus the commercially insured population using the national surveys of health care providers conducted by the National Center for Health Statistics.

A SYSTEM OF INDICATORS

The monitoring system will provide indicators of performance at the level of the individual rehabilitation facility, the rehabilitation market, the post-acute care market, and the Medicare program as a whole. The recommendations for the design and implementation of the monitoring system will include specific rules for each indicator--including the calculation formula(s), the source of each data element, and the level of the indicator (e.g., hospital versus MSA). For hospital statistics, we will include a method to determine if changes in the indicator are statistically significant.

We will also describe how the collection of indicators can be analyzed to produce a coherent description of changes that are occurring in post-acute care. We will tabulate (or estimate) each indicator for each year from 1996 through 2001. This will provide baseline measures as well as information on trends and year-to-year variability.

Our design for the monitoring system will not only describe the indicators, it will also show how they can be used together to obtain a clear description of access, outcomes, and cost at inpatient rehabilitation and other post-acute care facilities.

Ultimately, we hope to shed light on each of the broad questions outlined at the beginning of this section. The results of these inquiries will be fed back into our efforts to refine the IRF PPS and to understand the far-reaching effects of payment system changes. We hope that the system of indicators we develop will form the basis for a monitoring system that will give policymakers an understanding of how the system is performing on an ongoing basis and whether it is furthering the goals of the Medicare program.

- 335 -

REFERENCES

American Medical Association (1999). *AMA Graduate Medical Education Directory, 1998-1999*. Washington, DC: American Medical Association.

Breiman, L. J., M. Friedman, R. A. Olshen, and C. J. Stone (1984). *Classification and Regression Trees*. Belmont, CA: Wandsworth, Inc.

Buchanan, J. L., P. Andres, S. Haley, S. Paddock, D. C. Young, and A. Zaslavsky (forthcoming). *Final Report on Assessment Instruments for PPS*. Santa Monica, CA: RAND, MR-1501.

Carter, Grace M., Joan L. Buchanan, Tenzing Donyo, Moira Inkelas, and Karen L. Spritzer (1997). *A Prospective Payment System for Inpatient Rehabilitation*. Available from National Technical Information Service, Springfield, VA (www.ntis.gov). Publication no. PB98106024.

Carter, Grace M., and Donna O. Farley (1993). *Interaction of Outlier Payment Policy with DRG Refinement and Recalibration*. Santa Monica, CA: RAND, DRU-447-HCFA.

Carter, Grace M., J. P. Newhouse, and D. A. Relles (1991). *Has DRG Creep Crept Up?* Santa Monica, CA: RAND, R-4098-HCFA/ProPAC.

Carter, Grace M., Daniel A. Relles, and Joan L. Buchanan (1997). *A Classification System for Inpatient Rehabilitation Patients: A Review and Proposed Revisions to the Functional Independence Measure-Function Related Groups*. Available from National Technical Information Service, Springfield, VA (www.ntis.gov). Publication no. PB98105992.

Carter, Grace M., Daniel A. Relles, Barbara O. Wynn, Jennifer Kawata, Susan M. Paddock, Neeraj Sood, and Mark E. Totten (2000). *Interim Report on an Inpatient Rehabilitation Facility Prospective Payment System*. Santa Monica, CA: RAND, DRU-2309-HCFA (forthcoming as MR-1503).

Carter, Grace M., and J. A. Rogowski (1992). How Pricing Policies, Coding, and Recalibration Method Affect DRG Weights, *Health Care Financing Review*, Vol. 14, No. 2, Winter, pp. 83-96. A slightly expanded version is available from RAND as MR-156-HCFA (1993).

Carter, Grace M., and J. D. Rumpel (1993). *An Evaluation of Medicare Payments for Transfer Cases*. Santa Monica, CA: RAND, MR-304-HCFA.

Centers for Medicare and Medicaid Services (2001). Medicare Program; Prospective Payment System for Inpatient Rehabilitation Facilities; Final Rule, *Federal Register*, Vol. 66, No. 152, Tuesday, August 7, 2001, pp. 41315-41430.

Cohen, M. A., and A. Tumlinson (1997). Understanding the State Variation in Medicare Home Health Care: The Impact of Medicaid Program

Characteristics, State Policy and Provider Attributes, *Medical Care*, Vol. 35, No. 6, pp. 618-633.

Cotterill, P., J. Bobula, and R. Connerton (1986). Comparisons of Alternative Relative Weights for Diagnosis-Related Groups, *Health Care Financing Review*, Vol. 7, No. 3, Spring 1986, pp. 37-51.

Dalton, K., and E. C. Norton (2000). Revisiting Rogowski and Newhouse on the Indirect Cost of Teaching: A Note on Functional Form and Retransformation in Medicare's Payment Formulas, *Journal of Health Economics*, Vol. 19, No. 16, pp. 1027-1046.

Department of Health and Human Services (DHHS) (1991). Prospective Payment System for Inpatient Hospital Capital Related Costs; Final Rule, *Federal Register*, Vol. 56, No. 169, August 30, 1991, pp. 43358-43524.

Duan, Naihua (1983). Smearing Estimate: A Nonparametric Retransformation Method, *Journal of the American Statistical Association*, Vol. 78, No. 383, September, pp. 605-610.

Efron, B., and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman and Hall.

Gianfrancesco, F. D. (1990). The Fairness of the PPS Reimbursement Methodology, *Health Services Research*, Vol. 25, Part 1, April, pp. 1-23.

Hastie, T., and R. J. Tibshirani (1990). *Generalized Additive Models*. London: Chapman and Hall.

Hastie, T., R. J. Tibshirani, and J. Friedman (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag.

Health Care Financing Administration (2000). *Medicare Program; Prospective Payment System for Inpatient Rehabilitation Facilities*. Federal Register, Notice of Proposed Rule Making, November 3, 2000.

Hosek, S., R. Kane, M. Carney, et al. (1986). *Charges and Outcomes for Rehabilitative Care: Implications for the Prospective Payment System*. Santa Monica, CA: RAND, R-3424-HCFA.

Kane, R. L., Q. Chen, M. Finch, et al. (1997a). *Functional Outcomes of Post-Hospital Care for Stroke and Hip Fracture Patients Under Medicare*. Manuscript, University of Minnesota School of Public Health.

Kane, R. L., Q. Chen, M. Finch, et al. (1997b). *The Optimal Outcomes and Costs of Post-Hospital Care Under Medicare*. Manuscript, University of Minnesota School of Public Health.

Keeler, E. B., G. M. Carter, and S. Trude (1988). Insurance Aspects of DRG Outlier Payments, *Journal of Health Economics*, Vol. 7, pp. 193-214.

Keith, R. A., D. B. Wilson, and P. Gutierrez (1995). Acute and Subacute Rehabilitation for Stroke: A Comparison. *Archives of Physical Medicine and Rehabilitation*, Vol. 76, pp. 495-500.

Kenney, G. M., and L. C. Dubay (1992). Explaining Area Variation in the Use of Medicare Home Health Services, *Medical Care*, Vol. 30, No. 1, pp. 43-57.

Kramer, A. M., J. F. Steiner, R. E. Schlenker, T. B. Eilertsen, C. A. Hrincevich, D. A. Tropea, L. A. Ahmad, and D. G. Eckhoff (1997). Outcomes and Costs After Hip Fracture and Stroke: A Comparison of Rehabilitation Settings, *JAMA*, Vol. 277, No. 5, pp. 396-404.

Lave, Judith R. (1985). Is Compression Occurring in DRG Prices? *Inquiry*, Vol. 22, Summer, pp. 142-147.

Lee, A. J., J. H. Huber, and W. B. Stason (1997). Factors Contributing to Practice Variation in Post-Stroke Rehabilitation, *Health Services Research*, Vol. 32, No. 2, pp. 197-221.

Liu, K., D. Wissoker, and C. Rimes (1998). Determinants and Costs of Post Acute Care Use of Medicare SNFs and HHAs, *Inquiry*, Vol. 35, No. 1, pp. 49-61.

Liu, K., et al. (1999). *Medicare's Post Acute Care Benefit: Background, Trends, and Issues to Be Faced*. Washington, DC: The Urban Institute. Available at http://aspe.os.dhhs.gov/daltcp/reports/mpacb.htm.

Medicare Payment Advisory Commission (MedPAC) (2001). *Report to the Congress: Medicare Payment Policy*. Washington, DC: MedPAC.

Medicare Payment Advisory Commission (MedPAC) (1998). *Health Care Spending and the Medicare Program: A Data Book*. Washington, DC: MedPAC.

Munin, M. C., et al. (1998). Early Inpatient Rehabilitation After Elective Hip and Knee Arthroplasty, *Journal of the American Medical Association*, Vol. 279, No. 11, pp. 847-852.

Newhouse, Joseph P., Shan Cretin, and Christina J. Witsberger (1989). Predicting Hospital Accounting Costs, *Health Care Financing Review*, Vol. 11, No. 1, Fall, pp. 25-33.

O'Dougherty, S. M., P. G. Cotterill, S. Phillips, E. Richter, N. DeLew, B. Wynn, and T. Ault (1992). Medicare Prospective Payment Rates Without Separate Urban and Rural Rates, *Health Care Financing Review*, Vol. 14, No. 2, pp. 31-47.

Prospective Payment Assessment Commission (1994). *Report and Recommendations to the Congress*. Washington DC: Prospective Payment Assessment Commission.

Relles, D. A., and G. M. Carter (forthcoming). *Linking Medicare and Rehabilitation Hospital Records to Support Development of a*

Rehabilitation Hospital Prospective Payment System. Santa Monica, CA: RAND, MR-1502-CMS.

Rogowski, J., and P. Byrne (1990). Comparison and Evaluation of Alternative DRG Weight Recalibration Methods, *Health Care Financing Review*, Fall.

Rogowski, J., and J. Newhouse (1992). Estimating the Indirect Costs of Teaching, *Journal of Health Economics*, Vol. 11, pp. 153-173.

Stineman, M. G., J. J. Escarce, J. E. Goin, B. B. Hamilton, C. V. Granger, and S. V. Williams (1995). Efficiency Pattern Analysis for Medical Rehabilitation, *American Journal of Medical Quality*, Vol. 10, No. 4, pp. 190-198.

Stineman, M. G., J. E. Goin, B. B. Hamilton, and C. V. Granger (1994). A Case-Mix Classification System for Medical Rehabilitation, *Medical Care*, Vol. 32, No. 4, pp. 366–379.

Stineman, M. G., B. B. Hamilton, C. V. Granger, J. J. Escarce, J. E. Goin, and S. V. Williams (1994). Four Methods of Characterizing Disability in the Formation of Function Related Groups, *Archives of Physical Medicine*, Vol. 75, pp. 1277–1283.

Stineman, M. G., A. J. Jette, R. C. Fiedler, and C. V. Granger (June 1997a). Impairment-Specific Dimensions Within the Functional Independence Measure, *Archives of Physical Medical Rehabilitation*, Vol. 78, pp. 636-643.

Stineman, M. G., C. J. Tassoni, J. J. Escarce, J. E. Goin, C. V. Granger, R. C. Fiedler, and S. V. Williams (October 1997b). Development of Function-Related Groups, Version 2.0: A Classification System for Medical Rehabilitation, *Health Services Research*, Vol. 32, No. 4, pp. 529-548.

Thorpe, K. E., Shan Cretin, and E. B Keeler (1998). Are the Diagnosis-Related Group Case Weights Compressed? *Health Care Financing Review*, Vol. 10, No. 2, pp. 37-46.

Uniform Data System for Medical Rehabilitation (UDSmr) (1997). *Guide for the Uniform Data Set for Medical Rehabilitation, Version 5.1*. Buffalo, NY: UDSmr.

Weissman, J. P., P. Dryfoos, and K. London (1999). Income Levels of Bad-Debt and Free-Care Patients in Massachusetts Hospitals, *Health Affairs*, Vol. 18, No. 4, pp. 156-166.